UNIVERSITEIT GENT

FACULTEIT PSYCHOLOGIE EN PEDAGOGISCHE WETENSCHAPPEN

Automatic relational stimulus processing

Marten Niclas Heider

Promotor: Prof. Dr. Adriaan Spruyt
Co-promotor: Prof. Dr. Jan De Houwer

Proefschrift ingediend tot het behalen van de academische graad
van Doctor in de Psychologie

2015

ACKNOWLEDGEMENTS

First and foremost, I want to thank my supervisors *Adriaan Spruyt* and *Jan De Houwer*. I owe a great deal of gratitude to both of them. *Adriaan*, I immensely profited from your expertise and your guidance, your perfectionism and your focus on detail. You taught me the importance of scrutinizing my data and how to write in a (slightly more) understandable manner, and you corrected the same grammatical errors in my drafts with only getting mildly angry at me from time to time. Thank you for the enlightening, fast-paced discussions and letting me see the bigger picture in my research. I always left your office understanding something at least a little better than when entering it. *Jan*, I am deeply indebted to you for the opportunity of this doctorate. I could not be more thankful for the working environment you provided and your support throughout the years. Thank you for giving me the chance to proof myself as a researcher.

I also want to thank *Anne Gast*. As a member of my doctoral guidance committee, an integral part of the LIPlab, the co-supervisor of my diploma thesis, and as seminar lecturer in my undergraduate studies, you accompanied my entire academic career. You guided me in my first small steps into the world of academic research in 2006, and brought my attention to the possibility of applying for a doctorate in Gent. You were one of the people that made me think about doing research in the first place. Thank you for everything!

A big thanks to all the other people that contributed to this thesis. To the external members of my guidance committee, *Sarah Teige-Mocigemba* and *Dirk Hermans*: Thank you for your time and constructive comments on my work. To my colleagues, *Agnes*, *Ama*, *Baptist*, *Colin*, *Evelien*, *Evelyne*, *Gaëtan*, *Helen*, *James*, *Jolien*, *Kristina*, *Maarten*, *Marijke*, *Pieter*, *Riccardo*, *Sean*, *Senne*, *Tom*, and *Yang*: Thank you for your comments and your help. To my office mates, *Tom* and *Maarten*: Thank you guys for always listening to my problems and for your help throughout the years. I could not have wished for better office mates. I always enjoyed coming to work and see you guys there.

When I started my PhD, I also started a new phase of my life in an exotic, foreign country. This change was, at times, tougher than I expected beforehand. I therefore want to thank the people that welcomed me in Gent and made me feel at home. Thanks to *Kristina* for smoothening my transition by including me in her group of friends. Also, a big thanks to them: *Antonio*, *Igor*, *Ineke*, *Jasmina*, *Jonathan*, *Maud*, *Naomi*, *Selene*, *Sofie*, and *Valentina*, thank you for welcoming me with open arms and making my life apart from academia so enjoyable.

I also would like to thank my parents and my girlfriend. *Anne* and *Meinolf*, you have always supported me unconditionally in my decisions, and gave me the confidence to pursue the things I deemed worthy of pursuing. Thank you! *Juli*, you supported me in so many ways throughout the last four years, and you always made me enjoy life apart from academia. I could not be more grateful to have you in my life.

Niclas

June 2015

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

INTRODUCTION

People can behave in two differing and sometimes conflicting ways. In line with a layperson's understanding of behavior, people are on one side capable of behaving in a deliberate and reflective (i.e., non-automatic) manner. Under certain circumstances, however, people behave impulsively and in contrast to their intentions (i.e., automatically[1]).

For instance, people are in general capable of self-regulating and controlling their drinking behavior. The consumption of alcohol can, for instance, be inhibited by negative expectancies of alcohol use (i.e., that drinking will produce negative outcomes). Then again, people sometimes behave in contrast to their intentions not to drink excessively (e.g., when binge drinking; Wiers, Houben, Smulders, Conrod, & Jones, 2006). Such a conflict between automatic and non-automatic behavior can best be observed in the behavior of addicts, who continue to excessively consume alcohol despite being aware of the negative consequences. However, alcohol addiction is only one example for the conflict of non-automatic and automatic behavior. It has also been shown, for instance, that people often act habitually against their dieting behavior (Verplanken & Aarts, 1999) or demonstrate discriminative behavior against a member of another ethnical group despite a self-reported endorsement of egalitarian beliefs (e.g., Greenwald, Oakes, & Hoffman, 2003; Payne, 2001).

[1] Note, that the concept of automaticity will be elaborated on in detail below. For now, the layperson's idea of automaticity suffices.

These observations have triggered a surge of interest in mental processes that mediate behavior. At the moment, dual-process models constitute the predominant way of thinking not only about how behavior is guided, but also regarding reasoning (Sloman, 1996), attitude change (e.g., Gawronski & Bodenhausen, 2006; Petty & Cacioppo, 1986), stereotyping (e.g., Devine, 1989) and a multitude of other facets of human cognition (for a review, see Smith & DeCoster, 2000). As implied by their name, all dual-process models are unified in the general assumption of two modes of information processing.

In this introduction, I will elaborate on these two modes of processing, the systems they operate in, and the mechanisms by which they guide behavior. Further, I will present methods to experimentally assess automatic behavior, and the rationale by which they make inferences about the underlying processes. A general understanding of these methods is crucial for the understanding of the experimental research presented in this dissertation. Essentially, the starting point of my dissertational research is the idea that all behavior might be guided by only one type of mental processing, as proposed by De Houwer and colleagues (De Houwer, 2009, 2014; Mitchell, De Houwer, & Lovibond, 2009). Accordingly, a short overview of such a single-process model will be given. The introductory chapter concludes with an overview of the empirical chapters of my dissertation.

DUAL PROCESS MODELS

For some time now, cognitive psychologists have tried to explain the conflict between automatic and non-automatic behavior by proposing mechanisms by which behavior is causally mediated by mental processes and representations. As indicated above, *dual-process* models of cognition propose two modes of information processing that operate in distinct ways within two separate systems of representations. Depending on the specific nomenclature the names for these two processes and systems differ (cf. Smith & DeCoster, 2000, p. 111, Table 1). I will refer to the processes, based on their defining structural difference, as *associative* and *propositional* processes. The systems in which they operate are named, accordingly, the associative and propositional system. In general, these two systems differ in the way in which information is represented, stored, and processed. These differences determine the way in which resulting behavior is perceived. Specifically, whether a behavior is seen as automatic or non-automatic is a direct consequence of the processes that initiated the specific behavior. In general, a process can be called automatic in the sense that it can for instance operate unintentionally, uncontrollable, goal independent, purely stimulus driven, unconsciously, efficient, or fast (see Moors & De Houwer, 2006). However, these different features of automaticity do not always co-occur, and different processes can possess different features of automaticity (e.g., De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). Automaticity is not a distinct state but rather an umbrella term for multiple features processes can possess.[2] Importantly, all dual-process models postulate consistently that the associative system operates under conditions of automaticity, while the rule-based processing within the propositional system requires cognitive capacity and motivation (e.g., Smith & DeCoster, 2000). Accordingly, behavior that is guided by the associative system is perceived as occurring under conditions of automaticity, whereas behavior that is based on propositional processing is seen as being non-automatic. Note, that this notion, i.e., that the two processing modes are divided along the concept of automaticity, is crucial within the context of my research.

In the subsequent paragraphs the associative and propositional systems, their modes of processing, and the mechanisms by which they guide behavior will be discussed separately. My remarks will be closely tied to the reflective-impulsive model of Strack and Deutsch (2004), as behavior was integrated into their model of cognition in a most comprehensive manner.

The associative system

The associative system is thought of as a simple associative network. This idea originated from semantic network models of memory (e.g., Anderson, 1983; Collins & Loftus, 1975) according to which long term memory is represented as a network of

[2] Note, that this definition of automaticity applies to behavior to the same extent. Behavior can be automatic in the sense that it us unintentional, uncontrollable, independent of goals, stimulus driven, unconscious, efficient, or fast.

associations between mental concepts (or aspects of concepts)[3]. It is assumed that the associations between concepts are gradually formed and changed as a result of experience, based on temporal-spatial continuity and frequency of stimulus parings in the environment (e.g., Shanks, 2007). Consequentially, associations between concepts differ in their strength. Concepts are activated when their physical counterparts are perceived in the environment, or thought of (Strack & Deutsch, 2004). This activation then automatically spreads to other elements in proportion to the strength of the associative link and the strength of the initial activation of the concept. Elements within the associative system can be concepts, but also cognitive, affective or motor reactions (e.g., attention allocation, motivational tendencies, response programs; Strack & Deutsch, 2004). Consequentially, a behavior can be automatically initiated within the associative system by the activation of a concept that is associated with that behavior. Consider, for instance, the abovementioned example of habitual drinking behavior. A person, for whom the presence of a beer and its consumption co-occurred repeatedly in the past, has developed a strong association between the concept of beer and the behavioral scheme of drinking. Consequentially, the presence of a beer can automatically activate the associated behavior of drinking it (Strack & Deutsch, 2004). Indeed, research in the addiction literature indicates that for heavy drinkers the consumption of alcohol is linked to positivity or arousal (e.g., De Houwer, Crombez, Koster, & De Beul, 2004; Wiers, Van De Luitgaarden, Van Den Wildenberg, & Smulders, 2005; Wiers, Van Woerden, Smulders, & De Jong, 2002) or alcohol-related concepts (e.g., Friday night, feeling good; Stacy, Ames, & Grenard, 2006). The activation of such concepts can be seen as cues, that, once activated, automatically trigger action tendencies like the approach of alcohol (Palfai & Ostafin, 2003).

The propositional system

While automatic behavior is postulated to be guided by a system of mental associations, non-automatic behavior is assumed to be dependent on a system that

[3] For the sake of simplicity, this explanation will be restricted to localist models. Note, that distributed memory models (e.g., Masson, 1995), although structurally more complex, do not differ from localist models in the basic assumption that the activation of concepts is thought of in terms of the associative process of spreading activation.

consists of propositional representations (e.g., Gawronski & Bodenhausen, 2006; Strack & Deutsch, 2004). While the associative system can be thought of as a long-term memory system, the propositional system is assumed to possess the properties of a temporal storage, in that it generates declarative knowledge in the form of propositions. Structurally, propositions differ from associations in two crucial points. First, propositions consist of one or more elements that are related to one another by qualified links. These relations can be simple logical relations that define category membership (e.g., A *is* B or A *is not* B; Strack & Deutsch, 2004) but can also be more complex relations (e.g., A *implies* B or A *causes* B; Lagnado, Waldmann, Hagmayer, & Sloman, 2007). Relations convey specific meaning that exceeds meaning implied by purely associative connections between concepts. Take, for instance, the proposition “I am good”. It is represented by a *descriptive* relation between the concepts *self* (i.e., “I”) and *positivity* (i.e., “good”). The similar proposition “I want to be good”, in contrast, relates the same two concepts, but in terms of *desirability*. An associative link between the two concepts, in contrast, does not convey such specific meaning. The second structural difference between associations and propositions concerns their veracity. Propositions are statements about the world. Consequentially, they are (believed to be) either accurate or false (De Houwer, 2009, 2014). They, in short, possess a (subjective) truth value. Associations, in contrast, do not contain any information about their veracity (e.g., Strack & Deutsch, 2004). Take the two propositions from the example above. The specified relations in the propositions “I am good” and “I want to be good” imply a factuality of each statement that can be either (subjectively) true or false. In the propositional system, behavioral decisions are based on propositional processes. Specifically, reasoning processes lead to behavioral decisions that take into account aspects like the feasibility and the desirability of behaviors (Ajzen, 1991). Take, for instance, a person that developed an automatic tendency to drink alcohol when associated concepts are activated within the associative system. Propositional processing allows for a reflection and reasoning of these associations. Whether or not such propositional processes are capable of controlling the behavior depends on the availability of cognitive capacities necessary for the propositional system to function properly (e.g., Strack & Deutsch, 2004). The resulting awareness of the associations that lead to alcohol consumption could, for instance, lead to an avoidance of situations in which the person is confronted with stimuli that automatically elicit drinking behavior.

The measurement of propositional and associative processes

Cognitive processes cannot be observed directly. Rather, they can only be inferred through the observation of changes in behavior (e.g., De Houwer & Moors, 2012). Measurement procedures that assess propositional processing allow the respondent to behave in a deliberate and controlled manner. Paradigmatic measures are structured interviews and self-report questionnaires that contain rating scales (e.g., Likert scales, semantic differential scales). The respondent’s responses (i.e., answers) on such measures reflect propositional processes of reasoning, as cognitive capacities necessary for the propositional system to function properly are available to the respondent.

The assessment of associative processes in contrast takes place with measurement procedures that assess behavior under conditions of automaticity. Inferences about activated knowledge is made from response patterns in well-controlled computer tasks, often referred to as *implicit measures* (De Houwer, 2006; De Houwer et al., 2009). The term implicit measure is used both in relation to the assessment and the underlying effect (De Houwer, 2006; Fazio & Olson, 2003). When the term is used to refer to the measurement procedure, implicit is used as an equivalent for the term *indirect*, as the outcome of the measurement procedure is assumed to be an indirect reflection of the to-be-measured construct. When the term implicit measure is in contrast used to refer to the measurement outcome (i.e., the effect) of the measurement procedure, it reflects the fact that the to-be-measured construct is assessed under conditions of automaticity (e.g., De Houwer & Moors, 2012; De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009).

A multitude of measurement procedures that assess automatic processes has been developed in recent history. The most prominent tasks are the Affect Misattribution Procedure (AMP; Payne, Cheng, Govorun, & Stewart, 2005), the Go/No-go Association Test (gNAT; Nosek & Banaji, 2001), the semantic and evaluative priming task (Fazio, Jackson, Dunton, & Williams, 1995; Wittenbrink, Judd, & Park, 1997), and the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). All of these measurement procedures are based on the assumption that automatic behavior is guided by the activation of associatively connected mental representations in memory (for a review, see Hughes, Barnes-Holmes, & De Houwer, 2011). In the context of this dissertation, the IAT and the priming task are most important. Both tasks will therefore be discussed more deeply in the following paragraphs.

The IAT is a double categorization task that assesses the extent to which individuals associate certain concepts with each other. Imagine an IAT constructed to access the associative strength of alcohol and valence (i.e., automatic attitudes towards alcohol; e.g., Wiers et al., 2002). In such an IAT, the associative strength between two target concepts (i.e., alcoholic drinks vs. soft-drinks) and two attribute concepts (i.e., positive vs. negative valence) is assessed by presenting stimuli of these concepts one by one on a computer screen. On one series of trials, participants are asked to press one key whenever they see a stimulus that refers to an alcoholic drink (e.g., beer, wine) or a positive stimulus, and a second key whenever they see a stimulus that refers to a soft-drink (e.g., cola, juice) or a negative stimulus. In a second series of trials, the rule by which participants respond is reversed. Stimuli referring to alcoholic drinks and negative stimuli are assigned to the first key whereas stimuli referring to soft-drinks and positive stimuli are assigned to the second key. The difference in performance (i.e., speed and accuracy) between the one categorization condition (i.e., alcoholic drink / positive vs. soft-drink / negative) and the other categorization condition (i.e., alcoholic drink / negative vs. soft-drink / positive) is argued to reflect the strength of associations between concepts in memory (e.g., Greenwald et al., 1998). Consequentially, in the example above, faster responses in the first categorization condition relative to the second categorization condition indicate stronger associations between the concepts alcoholic drink and positivity (and soft-drink and negativity) as compared to alcoholic drink and negativity (and soft-drink and positivity). Such a difference in performance can be seen as reflective of a positive attitude towards alcohol, as strong associations in memory between a concept and valence are regarded to be representative of psychological attributes (i.e., attitudes, stereotypes; e.g., Greenwald et al., 2002). However, there is also evidence that differences in performance can reflect other attributes of the stimuli (e.g., salience, perceptual similarity) or the participant (e.g., cognitive skills; for an overview, see De

Houwer et al., 2009). In any case, regardless of how an observed difference in performance between the two conditions of the IAT is interpreted (e.g., being reflective of attitudes), the difference is argued to be based on the associative strength of the target concepts (e.g., alcoholic drinks) and the attribute concepts (e.g., positivity).

The priming task constitutes the second most popular implicit measure. It also allows for inferences of the associative strength between concepts. In typical priming tasks participants are presented, on each of a series of trials, with a prime stimulus followed by a target stimulus. Crucially, the relationship between the prime and the target is manipulated. On some trials both stimuli are somehow related (i.e., congruent trials), whereas both stimuli are unrelated on other trials (i.e., incongruent trials). The *priming effect* denotes the typical finding of superior performance (i.e., in speed and/or accuracy) in responding to target stimuli (e.g., pressing a button) on congruent trials as compared to incongruent trials. When participants in an *evaluative* priming task are asked to indicate whether targets denote positive or negative words, their response performance is influenced by the evaluative connotation of the prime stimuli (i.e., observable in the priming effect). Therefore, to assess the associative strength between alcohol and positivity or negativity with an evaluative priming task, prime stimuli would be either related or unrelated to the concept of alcohol (e.g., words or picture of well-known beer brands), while target stimuli would be positive or negative stimuli. Imagine a person performs better on trials consisting of alcohol related primes and positive targets relative to trials with alcohol related primes and negative targets. Such an observation allows for the inference that, for this person, the concept of alcohol is more closely related to the concept of positivity than to the concept of negativity. In terms of associative processing, the relation between the concept and its valence is thought of as an association between a concept and its valence (i.e., object-evaluation association; Fazio, Sanbonmatsu, Powell, & Kardes, 1986). In the words of Fazio (2001), the "presentation of an attitude object as a prime should activate any associated evaluation and, hence, facilitate a related judgment" (p. 116). Research has shown that this holds true for other object features apart from valence. What feature of the prime stimuli is processed depends on the aspect or feature participants attend their attention to (e.g., Spruyt, Houwer, & Hermans, 2009). Participants in an evaluative priming task are asked to categorize targets on the basis of their evaluative connotation. Accordingly, the evaluative connotation of the primes is processed. If, however, participants are asked to categorize targets on the basis of a semantic feature (e.g., animacity), this feature is processed. Crucially, regardless of what specific prime feature is processed, within dual-process models these instances of processing are thought of in terms of associations between the prime and the specific, task-relevant feature of the prime, which is automatically activated.

Conclusion

Dual-process models of cognition currently constitute the predominant way of explaining automatic and non-automatic behavior. They propose two modes of information processing to be responsible for the two types of behavior, respectively. Automatic behavior is assumed to be guided by associative processing, which is thought of in terms of an automatic activation of associations between concepts that are structured based on similarity and contiguity, and established and strengthened over many experiences. Accordingly, measurement procedures that observe behavior under conditions of automaticity are used for their assessment. Non-automatic behavior is in contrast assumed to be guided by propositional processes. Such processes allow for a deliberate reasoning about activated information. Crucially, the complexity of information in the propositional system differs from that stored in the associative system. Specifically, propositions allow for specified relations between concepts and the assignment of a (subjective) truth value. Propositions are, consequentially, statements about the world, as compared to simple associations between concepts. Their construction is assumed to be time-consuming and to depend upon cognitive resources (e.g., Strack & Deutsch, 2004). Their assessment therefore relies of measurement procedures that allow for non-automatic, deliberate responding (e.g., interviews, questionnaires).

AN ALTERNATIVE VIEW: THE PROPOSITIONAL SINGLE-PROCESS APPROACH

Purely propositional accounts of cognition (De Houwer, 2009, 2014; Mitchell et al., 2009) constitute a recent alternative to the predominant dual-process models. These single-process models dismiss the dichotomy of processing modes along the concept of automaticity in favor of the assumption that all behavior, both automatic and non-automatic, is guided by the propositional system. Accordingly, propositional processes are assumed to operate under automaticity conditions as well as non-automaticity conditions. The associative system is disregarded in favor of a single propositional system, in which knowledge is organized in the form of propositions. Both automatic and controlled processes contribute to the formation and activation of knowledge stored in this system.

As indicated above, dual-process models recognize the role of propositional processes in the explanation of non-automatic behavior. Accordingly, for a purely propositional model to gain credibility, it has to provide an explanation of automatic behavioral effects. The first part of the research presented in this dissertation focused on exactly that. Chapter 2 consists of an empirical investigation of the question whether propositional processes can influence behavior under automaticity conditions. Furthermore, a mechanism that can account for propositional processing under automaticity conditions is proposed.

The research in the second part of this dissertation in contrast focused on an implication of propositional single-process models. If automatic behavior is indeed guided by propositional processes, the scope of implicit measures should increase drastically. As indicated above, implicit measures can be used to infer how concepts are related to another. Assuming only one informational system that is structured by propositions, implicit measures might be sensitive to aspects that are specific to propositions. As argued above, propositions structurally differ from associations, in that they possess qualified links that specify the relation between concepts, and possess a truth value. Accordingly, it makes sense to examine the possibility of implicit measures to be capable of assessing (a) how concepts are related in memory and (b) whether or not the resulting proposition is believed to be true or false. Such an implicit measure should consequentially allow to assesses a person’s beliefs (i.e., propositions that are believed to be true), rather than simple associations between objects and their features (e.g., alcohol – bad). In the second part of this dissertation, which comprises of Chapters

3 to 5, research on the development and empirical validation of such an implicit measure of beliefs is presented.

Conclusion

Dual-process models propose stimulus processing under conditions of automaticity to be guided by associative processes. Propositional single-process models in contrast postulate that automatic behavior is guided by propositions. The empirical research presented in this dissertation addressed this discrepancy. Specifically, in one line of research it was examined whether propositional processes can influence behavior under automaticity conditions. A second line of research addressed an implication of the single-process propositional models, namely the increased capacities of implicit measures in assessing psychologically relevant constructs. Both research lines are outlined in more detail in the next section.

OVERVIEW OF THE EMPIRICAL CHAPTERS

Research line 1: The demonstration of automatic propositional processing

The focus in **Chapter 2** is directed at the question whether the processing of an object and its relation in size to another object can occur under conditions of automaticity. As information about the relation of concepts is unique in propositional knowledge, the process of constructing knowledge that contains such information can be considered an instance of propositional processing (i.e., propositional categorization; Strack & Deutsch, 2004). Crucially, dual-process models postulate that such a propositional categorization cannot operate under conditions of automaticity, as automatic processes are equated with associative processes. A series of four sequential priming studies on the automatic processing of relational information about size (i.e., “smaller than”, “larger than”) was conducted. In each study, participants were presented on each trial with two line drawings that were similar in size but depicted objects of different size (e.g., a bike and a house). Participants were asked to judge the size of the second object (i.e., the target) relative to the size of a particular reference object. Participants, for instance, were asked to decide whether a house is smaller or larger than a football. The size of the reference object varied between blocks of trials or from trial to trial, being either a football or a car. If it changed from trial to trial, it was presented at the offset of each trial. The line drawing that preceded the target stimulus (i.e., the prime) depicted an object that was also smaller or larger than the reference object. Trials in such a setup can be grouped into two trial conditions. On size-congruent trials, the size of the prime and the target stimuli relative to the reference object were the same (e.g., both stimuli were smaller than the reference object). On size-incongruent trials, in contrast, the relation in size of prime and target stimuli relative to the reference object was different (e.g., the prime was smaller and the target was larger than the reference object). Task performance was expected to be dependent upon the size congruency of prime and target. Specifically, faster and more accurate responses to targets were expected on size-congruent trials as compared to on size-incongruent trials. To account for the task-specific processing of the prime that is necessary for such priming effects, the mechanism of *task misapplication* was proposed. Participants were expected to unconsciously misapply the target-task (i.e., the judgment of the size of the targets relative to the reference object of the trial) to the primes. Consequentially, the processing of the primes was expected to lead to the generation of propositional information about the size of the prime relative to the size of the reference object of the trial (e.g., “the football is smaller than the car”). Crucially, primes differed in their size relative to the size of the two reference objects (e.g., the prime object bike is smaller than the car but larger than the football). Accordingly, whether a certain prime-target pair belonged to the group of size-congruent or size-incongruent trials was dependent on the size of the reference object of the trial. The two objects bike and house are, for instance, size-congruent relative to the football (i.e., both are larger) but size-incongruent relative to the car (i.e., the bike is smaller, but the house larger than a car). In general, primes were expected to facilitate “larger than” responses when they were bigger than the reference object on that trial (i.e., the football) and “smaller than” responses when they were smaller than the reference object on that trial (i.e., the car). Thus, priming effects that were dependent on the size of the reference object were expected to occur.

As expected, such *relational priming effects* were observed in all four experiments. The finding of these behavioral effects implies, as argued above, that propositional knowledge about the size of the task-irrelevant prime stimuli in relation to the reference object was automatically generated on each trial. As discussed in detail in Chapter 2, the observation of relational priming effects can be seen as empirical evidence of the occurrence of propositional processes under conditions of automaticity. Such findings are difficult to reconcile with dual-process models. They, instead, fit nicely to the idea that propositional processes can occur under conditions of automaticity, as proposed by propositional single-process accounts of cognition. Moreover, the empirical evidence that propositional processes can be executed under conditions of automaticity constituted the basis of the second research line of my dissertation.

Research line 2: The indirect assessment of beliefs

If automatic behavior can be guided by propositional processes, the scope of implicit measures should increase drastically in that implicit measures could be sensitive the truth value that is assigned to propositions. The findings presented in Chapter 2 support this assumption. Accordingly, the second part of this dissertation, containing Chapters 3 to 5, is devoted to the development and empirical validation of an implicit measure that is sensible to the truth value of propositional knowledge. Such a measure could be seen as an implicit measure of beliefs (i.e., propositions that are believed to be true). Throughout the three chapters that are dedicated to this endeavor, the concept of body dissatisfaction is of central importance. This construct was chosen for both practical and theoretical reasons. For once, the construct of body dissatisfaction constitutes a highly relevant psychological construct for young female adults, who make up the majority of available participants. Moreover however, body dissatisfaction can be defined as the negative attitude towards one’s own body resulting from a (perceived) discrepancy between one’s actual physical appearance (i.e., the actual body image) and internalized ideals about one’s physical appearance (i.e., the ideal body image; e.g., Cash & Szymanski, 1995). Therefore, the construct of body dissatisfaction comprises more than a simple association between the self and body size. Both the belief about the ideal and actual body image involve a relation between the concepts self and body size but, crucially, differ with regard to how the concepts are related. More specifically, beliefs about actual body image are characterized by a descriptive relation (e.g., I *am* thin). In contrast, beliefs about ideal body image relate the self to body size in terms of desirability (e.g., I *want to be* thin). Consequentially, body dissatisfaction can be seen as a function of two similar but distinct propositional beliefs. The separate assessment of these two beliefs can be considered as a test of whether propositions used in implicit measure of beliefs are indeed processed completely, that is, including the relational information.

First, however, in **Chapter 3** the central hypothesis that body dissatisfaction is dependent on implicit beliefs concerning the actual and ideal body image is investigated. Implicit beliefs about the actual and ideal body image were assessed with two versions of the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2006). The IRAP is a computerized reaction time task in which participants are presented with combinations of stimuli that are either in line or in contrast to a specific belief. Crucially, in one block of trials, participants are asked to respond as if they personally endorse that belief (e.g., a thin ideal body image). They are asked to select the response ‘true’ whenever a stimulus combination is presented that is in line with that specific belief (e.g., ‘I want to be + thin’), and to select the response ‘false’ whenever a stimulus combination is presented that is in contrast to that specific belief (e.g., ‘I don’t want to be + thin’). In a second block of trials, participants are asked to respond as if they do not endorse the belief in question. Accordingly, they are expected to select the response ‘true’ whenever a stimulus combination is presented is in contrast with the belief, and to select the response ‘false’ whenever a stimulus combination is presented that is in line with the belief. The IRAP is based on the assumption that it is more difficult to respond in a manner that is inconsistent with one’s personal beliefs than it is to respond in a manner that is consistent with one’s personal beliefs. IRAP scores can thus be used as an indicator of the degree to which a person endorses or rejects a certain belief. The beliefs concerning actual and ideal body image were assessed with two versions of the IRAP. The scores of the two IRAPs should reflect whether a participant is satisfied or dissatisfied with her body. Specifically, the implicit belief to be thin, assessed with an IRAP focusing on the actual body image, was expected to be more pronounced in participants low in body dissatisfaction as compared to participants high in body dissatisfaction. In contrast, the implicit desire to be thin, assessed with an IRAP focusing on the ideal body image, was expected to be less pronounced in participants low in body dissatisfaction as compared to participants high in body dissatisfaction. As will become clear, these general hypotheses were confirmed by the data. Furthermore, results indicated that both IRAP measures were related to body dissatisfaction and other variables (e.g., objective body size measures, explicit measures of actual and ideal body image) to a different extent. The findings indicated that for the assessment of body dissatisfaction it is key to examine not only whether the two concepts self and body size are related in memory but rather how they are related.

In **Chapter 4** the Relational Responding Task (RRT) is introduced as an alternative to the IRAP in capturing beliefs at the implicit level. Like in the IRAP, participants in an RRT are asked to respond in line with specific beliefs in different blocks of trials. In its initial test, the RRT was used to capture the extent to which Flemish participants held implicit racist beliefs about immigrants. Specifically, participants were asked to respond to statements that were either in line with the belief that Flemish people are more intelligent that immigrants (e.g., “Flemish people are smarter than immigrants”, “Immigrants are dumber than Flemish people”) or in line with the belief that immigrants are more intelligent than Flemish people (e.g. “Flemish people are dumber than immigrants”, “Immigrants are smarter than Flemish people”). In one block of trials, participants were asked to respond to these statements as if they believed that Flemish people are more intelligent than immigrants (i.e., to respond “true” to statements such as “Flemish people are smarter than immigrants” and to respond “not true” to statements such as “Flemish people are dumber than immigrants”). In a second block of trials, participants were asked to respond as if they believed that immigrants are more intelligent than Flemish people (i.e., to respond “not true” to statements like “Flemish people are smarter than immigrants” and to respond “true” to statements like “Flemish people are dumber than immigrants”). The RRT is based on the same assumption as the IRAP, i.e., that it is more difficult for participants to respond in a manner that is inconsistent with their personal beliefs than it is to respond in a manner that is consistent with their personal beliefs. This assumption is based on the idea that the processing of a statement leads to an automatic activation of its truth value. Consequentially, the scores of this RRT can be used as an indicator of the degree to which a person personally endorses or rejects the belief that Flemish people are more intelligent than immigrants. The RRT scores were expected to correlate with an explicit measure of the extent to which participants believed that Flemish people are more or less intelligent than immigrants. This hypothesis was confirmed. The RRT scores indeed indicated the strength of racist beliefs of the participants.

In **Chapter 5** the investigation of body dissatisfaction was revisited. This time, beliefs concerning actual and ideal body image were assessed with two RRTs. The dissociation between actual and ideal body image beliefs can be considered crucial in the validation process of the RRT, as it allows for testing the potential of the RRT to capture differences between beliefs that vary only in their relational component. Statements used in the RRT that addressed beliefs about the actual body image referred to how the own body is perceived (e.g., 'I possess a slim body', 'I weight too much'). Statements of the RRT that addressed beliefs about the ideal body image focused on an ideal state (e.g., 'I wish I possessed a slimmer body', 'I strive to weigh more'). Crucially, the statements used in the two RRTs differed from another in that the relevant concepts of self and body size were connected in a descriptive relation or a relation of desirability, respectively. Similar to the findings regarding implicit beliefs of actual and ideal body image with two IRAPs (in Chapter 3), participants were hypothesized to differ in their RRTs as a function of their level of explicitly measured body dissatisfaction. Specifically, the belief to be thin was expected to be more pronounced in participants low in body dissatisfaction as compared to participants high in body dissatisfaction. The desire to be thin, in contrast, was expected to be less pronounced in participants low in body dissatisfaction as compared to participants high in body dissatisfaction. As will be shown, these hypotheses were confirmed by the data. Results indicated that both RRT measures were related to body dissatisfaction and other variables (e.g., objective body size measures, explicit measures of actual and ideal body image) to a different extent. Furthermore, the prediction of body dissatisfaction was shown to benefit from RRT scores over and above explicit measures of body size.

Summarized, the focus in the second research line of this dissertation lies on the development and validation of an implicit measure that, based on the assumptions of propositional processing models, is assumed to allow for an assessment of beliefs rather than simple associations. In Chapter 3 it is investigated whether the construct of body dissatisfaction is appropriate for this endeavor. In Chapters 4 and 5 it is then focused on the development and the validation of the propositional implicit measure. The last two chapters both focus on substantiating the claim that the RRT can be used as a tool to capture beliefs at the implicit level.

CONCLUDING NOTE

To avoid confusion, the reader should note that each of the empirical chapters of this dissertation is based on an autonomous research paper that either has been submitted for publication or has been published. As each of the manuscripts should be able to stand on its own, the text of some of the chapters may partially overlap.

REFERENCES

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179–211. doi: 10.1016/0749-5978(91)90020-T

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*, 261–295. doi: 10.1016/S0022-5371(83)90201-3

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, *32*, 169–177.

Cash, T. F., & Szymanski, M. L. (1995). The Development and Validation of the Body-Image Ideals Questionnaire. *Journal of Personality Assessment*, *64*, 466–477. doi: 10.1207/s15327752jpa6403_6

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428. doi: 10.1037/0033-295X.82.6.407

De Houwer, J. (2006). What are implicit measures and why are we using them. In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of implicit cognition and addiction* (pp. 11–28). Thousand Oaks, CA: Sage.

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*, 1–20. doi: 10.3758/LB.37.1.1

De Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass*, *8*, 342–353. doi: 10.1111/spc3.12111

De Houwer, J., Crombez, G., Koster, E. H., & De Beul, N. (2004). Implicit alcohol-related cognitions in a clinical sample of heavy drinkers. *Journal of Behavior Therapy and Experimental Psychiatry*, *35*, 275–286. doi: 10.1016/j.jbtep.2004.05.001

De Houwer, J., & Moors, A. (2012). How to define and examine implicit processes. In R. Proctor & J. Capaldi (Eds.), *Psychology of science: Implicit and explicit processes* (pp. 183–198). Oxford: Oxford University Press.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368. doi: 10.1037/a0014211

Devine, P. G. (1989). Stereotypes and prejudice: their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18. doi: 10.1037/0022-3514.56.1.5

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition & Emotion*, *15*, 115–141. doi: 10.1080/02699930125908

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027. doi: 10.1037/0022-3514.69.6.1013

Fazio, R. H., & Olson, M. A. (2003). Implicit Measures in Social Cognition Research: Their Meaning and Use. *Annual Review of Psychology*, *54*, 297–327. doi: 10.1146/annurev.psych.54.101601.145225

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238. doi: 10.1037/0022-3514.50.2.229

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692–731. doi: 10.1037/0033-2909.132.5.692

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*, 3–25. doi: 10.1037/0033-295X.109.1.3

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Greenwald, A. G., Oakes, M. A., & Hoffman, H. G. (2003). Targets of discrimination: Effects of race on responses to weapons holders. *Journal of Experimental Social Psychology*, *39*, 399–405. doi: 10.1016/S0022-1031(03)00020-9

Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorizing in implicit attitude research: Propositional and behavioral alternatives. *The Psychological Record*, *61*, 464–498.

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation. *Causal Learning: Psychology, Philosophy, and Computation*, 154–172. doi: 10.1093/acprof:oso/9780195176803.003.0011

Masson, M. E. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 3–23. doi: 10.1037/0278-7393.21.1.3

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*, 183–198. doi: 10.1017/S0140525X09000855

Moors, A., & De Houwer, J. (2006). Automaticity: A Theoretical and Conceptual Analysis. *Psychological Bulletin*, *132*, 297–326. doi: 10.1037/0033-2909.132.2.297

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition*, *19*, 625–666. doi: 10.1521/soco.19.6.625.20886

Palfai, T. P., & Ostafin, B. D. (2003). Alcohol-related motivational tendencies in hazardous drinkers: assessing implicit response tendencies using the modified-IAT. *Behaviour Research and Therapy*, *41*, 1149–1162. doi: 10.1016/S0005-7967(03)00018-4

Payne, B. K. (2001). Prejudice and perception: the role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, *81*, 181–192. doi: 10.1037/0022-3514.81.2.181

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*, 277–293. doi: 10.1037/0022-3514.89.3.277

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123–205. doi: 10.1016/S0065-2601(08)60214-2

Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *The Quarterly Journal of Experimental Psychology*, *60*, 291–309. doi: 10.1080/17470210601000581

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22. doi: 10.1037/0033-2909.119.1.3

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108–131. doi: 10.1207/S15327957PSPR0402_01

Spruyt, A., Houwer, J. D., & Hermans, D. (2009). Modulation of automatic semantic priming by feature-specific attention allocation. *Journal of Memory and Language*, *61*, 37–54. doi: 10.1016/j.jml.2009.03.004

Stacy, A. W., Ames, S. L., & Grenard, J. (2006). Word association tests of associative memory and implicit processes: Theoretical and assessment issues. In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of implicit cognition and addiction* (pp. 75–90). Thousand Oaks, CA: Sage.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247. doi: 10.1207/s15327957pspr0803_1

Verplanken, B., & Aarts, H. (1999). Habit, attitude, and planned behaviour: is habit an empty construct or an interesting case of goal-directed automaticity? *European Review of Social Psychology*, *10*, 101–134. doi: 10.1080/14792779943000035

Wiers, R. W., Houben, K., Smulders, F. T., Conrod, P. J., & Jones, B. T. (2006). To drink or not to drink: the role of automatic and controlled cognitive processes in the etiology of alcohol-related problems. In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of implicit cognition and addiction* (pp. 339–361). Thousand Oaks, CA: Sage.

Wiers, R. W., Van De Luitgaarden, J., Van Den Wildenberg, E., & Smulders, F. T. (2005). Challenging implicit and explicit alcohol-related cognitions in young heavy drinkers. *Addiction*, *100*, 806–819. doi: 10.1111/j.1360-0443.2005.01064.x

Wiers, R. W., Van Woerden, N., Smulders, F. T. Y., & De Jong, P. J. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology*, *111*, 648–658. doi: 10.1037/0021-843X.111.4.648

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, *72*, 262–274. doi: 10.1037/0022-3514.72.2.262

CHAPTER 2

ON THE AUTOMATICITY OF RELATIONAL STIMULUS PROCESSING [1]

ABSTRACT

While it is widely accepted that the semantic analysis of a stimulus can take place in an automatic fashion, it is typically assumed that non-automatic processes are required to process the relation of one stimulus relative to other stimuli. Nevertheless, there is evidence to support the idea that such relational stimulus processing can also take place under automaticity conditions. We examined this hypothesis further in four sequential priming experiments in which participants were asked to categorize target objects as larger or smaller than a reference object (i.e., a football or a car). Crucially, some primes were objects that were larger than the small reference object but smaller than the large reference object (e.g., a bike). Results showed that the impact of these primes upon target responding was dependent on the size of the reference object. When the size of the reference object was small, these primes facilitated responses towards large targets relative to small targets. Vice versa, when the size of the reference object was large, the same set of primes facilitated responses towards small targets relative to large targets. This result was obtained when the size of the reference object was manipulated block-wise (Experiments 1 & 3), trial-wise (Experiments 2 & 4), and even when the primes were presented near subjective recognition thresholds (Experiment 4). Taken together, our findings provide strong evidence for the hypothesis that complex relational stimulus processing can take place under automaticity conditions. A possible underlying mechanism is proposed.

[1] Based on Heider, N., Spruyt, A., & De Houwer, J. (2015). On the automaticity of relational stimulus processing. *Manuscript submitted for publication.*

INTRODUCTION

There is little doubt that the semantic analysis of a stimulus can take place in an automatic fashion (e.g., Bargh, Schwader, Hailey, Dyer, & Boothby, 2012; Nosek, Hawkins, & Frazier, 2011). It is well-established, for example, that one can process the evaluative tone of a stimulus in a fraction of a second (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986), even in the absence of plenteous cognitive resources (e.g., Hermans, Crombez, & Eelen, 2000; but see Klauer & Teige-Mocigemba, 2007), an explicit evaluative processing mindset (e.g., Spruyt, Hermans, Houwer, & Eelen, 2002), or the conscious identification of this stimulus (e.g., Draine & Greenwald, 1998). Although these findings do not imply that the evaluation of a stimulus is unconditional (i.e., automatic in all possible ways; see Moors & De Houwer, 2006), it is safe to conclude that it has several of the features that fall under the umbrella of the concept automaticity. Likewise, there is ample evidence showing that automatic stimulus processing plays a key role in domains such as face perception and social judgment (e.g., Bargh et al., 2012), consumer behavior (e.g., Dijksterhuis, Smith, van Baaren, & Wigboldus, 2005), and health-related behavior (e.g., Wiers, Van Woerden, Smulders, & De Jong, 2002).

All these observations connect well with the idea that the ability to automatically process incoming stimulus information facilitates our interaction with the environment in important ways. In many situations, however, adaptive behavior requires a comparative integration of various sources of information rather than an assessment of a single stimulus. As an example, consider the simple act of buying a box of Belgian chocolates. Imagine that you just selected one box of chocolates out of many alternatives and that you are comparing this box with all the other alternatives in terms of their price, quantity, quality, or taste. Research in social economics shows that such choices are highly dependent on the alternatives given in a specific context. For example, when presented with three similar chocolate boxes with increasing size and price (i.e., chocolate box A, B, and C, respectively), most consumers will prefer chocolate box B over chocolate box C (Simonson, 1989). If, however, chocolate box A is replaced by a chocolate box D that is larger and more expensive than chocolate box C, consumers are more likely to select chocolate box C than chocolate box B. More generally, in daily life, we are bombarded with situations in which it is important to take into account the relations between stimuli. The question therefore arises to what extent such relational information can be processed in an automatic fashion (i.e., fast, unintentional, uncontrolled, efficient, or unconscious, see Moors & De Houwer, 2006).

To shed light on this issue, we conducted a series experiments to examine whether the online comparison of two stimuli can take place under automaticity conditions. This research was partly inspired by and designed to extend earlier work of Dehaene et al. (1998) who convincingly showed that masked numerical primes can trigger cognitive process without gaining access to consciousness. Using an adaptation of the sequential priming paradigm, they presented numerals that were either smaller or larger than five as (masked) primes and (clearly visible) targets. Crucially, participants were asked to decide as quickly as possible whether the target numerals were larger or smaller than five. Results clearly showed that performance in this comparative judgment task was facilitated when the primes and the targets had the same relation to the reference number (e.g., 4 – 3, 6 – 9) than when their relation to the reference number was different (e.g., 4 – 9, 9 – 1). To explain this priming effect, the authors argued that participants misapplied the task instructions to the primes (Dehaene et al., 1998, p. 598). Moreover, if it is assumed that the use of masked primes is sufficient to rule out the operation of non-automatic processes (see Dehaene et al., 1998; Kiefer, 2002; Kiefer & Martens, 2010; Ulrich, Hoenig, Grön, & Kiefer, 2013; but see Moors & De Houwer, 2006), the findings of Dehaene et al. (1998) seem to suggest that the processing of the (size) relation between two stimuli can occur in an automatic fashion, a phenomenon that we will refer to as (an instance of) *automatic relational processing*.[2]

[2] One could argue that even the processing of a single stimulus is often relational, for instance, when determining the semantic category of that stimulus (i.e., relating the stimulus to a semantic category). In this regard, (subliminal) priming effects have repeatedly been observed on the basis of category membership (see Quinn & Kinoshita, 2008, for a discussion). We do not consider stimulus processing in these cases to be truly relational, as the observed priming effects are attributed to the processing of certain properties of one stimulus alone (i.e., the task-relevant category membership of the prime stimulus). In contrast, the definition of relational processing in this paper is restricted to the online comparison of the task-relevant features of two distinct stimuli.

It must be emphasized, however, that there are also reasons to doubt the scope and theoretical interpretation of the observations reported by Dehaene et al. (1998). First, the stimulus set used by Dehaene et al. (1998) was very small (i.e., four numerals smaller than 5 and four numerals larger than 5). Second, only a single reference stimulus was used throughout the experiment (i.e., the numeral 5). Third, the numerals that were used as primes were also used as targets. For these reasons, one might argue that the numerals became linked to either the response *smaller than* or *larger than* as participants worked through the experimental procedure. Accordingly, the priming effects obtained by Dehaene et al. (1998) may have been due to a direct activation of these stimulus-response links rather than an online processing of the size relation between the primes and the reference number five (see Ansorge, Kiefer, Khalid, Grassl, & König, 2010; Damian, 2001).

Some of the limitations of the work of Dehaene et al. (1998) were addressed in subsequent studies (Kunde, Kiesel, & Hoffmann, 2003; Naccache & Dehaene, 2001; Quinn & Kinoshita, 2008; Reynvoet, Gevers, & Caessens, 2005; but see Footnote 2). Most importantly, Van Opstal, Gevers, Osman, and Verguts (2010, Experiment 2) conducted a subliminal priming study in which participants were presented with prime and target compounds, each consisting of two stimuli. Participants were asked to decide whether the two target stimuli (i.e., two single digits) were the same. Results showed that the speed of these similarity judgments was dependent upon the nature of the preceding prime compound. More specifically, target responding was faster when the relationship between the elements of a target compound was identical to the relationship between the elements of a prime compound. For instance, the correct categorization of two identical target digits (e.g., 3 – 3) was faster when these target digits were preceded by a prime compound containing two equivalent letters (e.g., A – a) than when the prime compound consisted of two primes with different identities (e.g., A – d). As argued by Van Opstal et al. (2010), these observations are quite difficult to account for in terms of the activation of stimulus-response links or mere perceptual similarity. Instead, in line with Dehaene et al. (1998), they argued that participants unintentionally misapplied the target task to the prime compounds. Moreover, given that the prime compounds were masked and presented for a very brief period of time (i.e., 33 ms), the findings of Van Opstal et al. (2010) again demonstrate that relational stimulus processing can take place in the absence of stimulus awareness.

Nevertheless, one could still question whether relational processing was truly involved in the effects reported by Van Opstal et al. (2010). In their crucial second experiment (which was designed to minimize the impact of perceptual similarity), the prime compounds always consisted of an uppercase and a lowercase letter. Importantly, the equivalence relation between upper and lower case versions of the same letter (e.g., A = a) must be over-learned in literate subjects. It could thus be argued that their effects were driven by a retrieval of prior knowledge about the equivalence of the uppercase and lowercase letters rather than by an active, online comparison of the two elements of the prime compound.

This concern, however, is less likely to hold for the evidence reported by Musch and Klauer (2003). In their Experiments 5 – 8, participants were asked to judge the similarity of briefly presented prime words (i.e., 57 ms) and clearly visible target words in terms of a non-evaluative stimulus dimension (e.g., font color). Results showed that these similarity decisions were affected by the evaluative match of the primes and the targets. Whereas *match* responses were facilitated by an evaluative match and inhibited by an evaluative mismatch, *mismatch* responses were facilitated by an evaluative mismatch and inhibited by an evaluative match. This data pattern clearly shows that participants processed the evaluative relationship between the primes and the targets even though this was not required (or even counterproductive) for the task at hand. Accordingly, if it is assumed that participants engaged in the evaluative comparison process despite the absence of the intention to so, the observations by Musch and Klauer (2003) add further weight to the idea that relational processing can occur automatically (Moors & De Houwer, 2006).

Finally, the same reasoning applies for the studies reported by Moors and colleagues (Moors & De Houwer, 2001; Moors, De Houwer, & Eelen, 2004; Moors, De Houwer, Hermans, & Eelen, 2005). In their evaluative priming studies, participants were presented with intrinsically neutral prime stimuli that signaled whether performance in a preceding goal-inducing task had been successful. Results showed that primes indicating success facilitated positive target responses whereas primes indicating failure facilitated negative target responses. Crucially, in order to extract the motivational valence of the primes, participants were required to compare the primes with the desired end-state of the goal-inducing task, even though this was not required for the task of evaluating the targets and even though primes were presented only briefly (i.e., 200 ms). Therefore, the findings of Moors and colleagues can also been seen as a demonstration of automatic relational processing, at least in the sense of the automaticity features ‘fast’, ‘unintentional’, and ‘efficient’.

In sum, the literature strongly suggests that relational stimulus processing can proceed under various automaticity conditions. However, with the exception of the priming effects in size-judgment tasks discussed earlier (Dehaene et al., 1998; Naccache & Dehaene, 2001; Quinn & Kinoshita, 2008; but see Footnote 2), evidence for automatic relational stimulus processing is limited to the mere detection of the presence or absence of an equivalence relation. This is an important limitation because equivalence is just one of many ways in which stimuli can be related (see Hayes, Barnes-Holmes, & Roche, 2002). In addition, the few studies that did concern comparative relational information (e.g., “smaller than” or “larger than” in the studies by Dehaene et al., 1998) were inconclusive with regard to the relational nature of the observed effects (Kunde et al., 2003; Quinn & Kinoshita, 2008). Hence, the important question of whether automatic relational processing might occur for complex relational information still needs to be answered.

To shed light on this issue, we conducted four sequential priming studies in which line drawings of objects of various sizes were presented as primes and targets. Participants were asked to judge the size of the target objects relative to either a small reference object (i.e., a football) or a large reference object (i.e., a car). In line with earlier work (e.g., Dehaene et al., 1998; see also Ansorge, Kunde, & Kiefer, 2014; Kiefer & Martens, 2010), we assumed that participants would misapply the target task to the prime set. Given that the target task required a comparison between the targets and the reference objects, a misapplication of the target task to the prime set would thus result in an (automatic) comparison between the primes and the reference objects. In that sense, our studies go beyond earlier work showing that non-relational automatic stimulus processing (i.e., the analysis of a stimulus irrespective of its relation to other stimuli) is constrained by top-down regulation (e.g., Ansorge & Heumann, 2003; Ansorge, Heumann, & Scharlau, 2002; Ansorge et al., 2010; Ansorge, Kiss, & Eimer, 2009; Ansorge et al., 2014; Ansorge & Neumann, 2005; Kiefer & Martens, 2010).

Crucially, the prime set included medium-sized objects (e.g., a bike) that were larger than the small reference object and smaller than the large reference object (see Figure 1). We thus expected medium-sized prime objects (e.g., a bike) to be judged as relatively large when the reference object was small (i.e., the football) but as relatively small when the reference object was large (i.e., the car). In addition, we expected that the misapplication of the target task to the prime set would result in an activation of task-relevant responses (e.g., Ansorge et al., 2010; Wentura & Frings, 2008). Medium-sized prime objects were thus expected to facilitate target responding when they had the same size relation with the reference object (i.e., both 'larger than' or 'smaller than') as compared to when they had a different size relation with the reference object (i.e., 'larger than' the reference object whereas the target was 'smaller than' the reference object, or vice versa). For example, after the presentation of a football as a reference object and a bike as a (medium-sized) prime, we expected task performance to be better when the target was an airplane as compared to when the target was a spoon. Conversely, after the presentation of a car as a reference object, we expected the presentation of the exact same prime stimulus (i.e., the bike) to lead to better performance when the target was a spoon as compared to when the target was an airplane. We will refer to this effect as the *relational priming effect* (RPE).

It may be noted that the emergence of this RPE requires two sequential processes (Spruyt, Gast, & Moors, 2011). In a first step, participants have to engage in an (automatic) analysis of the relational information conveyed by the primes. It is this process that is at the heart of our research. In a second step, the outcome of this comparison process must be translated in an observable priming effect by means of some mechanism. There is an abundant number of studies focusing on these mechanisms, many of which are very likely to play a critical role in the emergence of the RPE in our experiments. As explained above, for example, we clearly capitalize on the well-known mechanism of direct response activation (e.g., Ansorge et al., 2010; Kiefer, Sim, & Wentura, 2015). It must be emphasized, however, that the experiments reported here were designed exclusively to shed light on the first process (i.e., automatic relational stimulus processing). The nature of the second process (e.g., response activation, semantic activation, etc.) is tangential to this research question and will therefore not be elaborated upon any further.

The aim of Experiment 1 was to establish the RPE, a cognitive marker of relational stimulus processing, under typical automaticity conditions. More specifically, the primes were presented for 200 ms and the time interval between the start of the presentation of the primes and the start of the presentation of the targets (i.e., the Stimulus Onset Asynchrony, SOA) was short (i.e., 250 ms). In addition, the stimuli that were used as primes were never presented as targets for an individual participant and participants were asked to respond to the targets only. Accordingly, should the RPE emerge under these conditions, it can be concluded that it is driven by fast-acting, unintentional processes (Moors & De Houwer, 2006; Moors, Spruyt, & De Houwer, 2010). After this initial demonstration of the RPE, Experiment 2 was performed to rule out two alternative interpretations. Next, building further on the work of Dehaene et al. (1998) and Van Opstal et al. (2010), Experiment 3 was designed to test the hypothesis that RPE arises as the result of the misapplication of the target task to the prime stimuli. Finally, in Experiment 4, we examined whether automatic relational processing can occur even in the absence of the conscious identification of the instigating stimulus.

EXPERIMENT 1

Method

Participants. Sixty-eight students at Ghent University (55 women, 13 men; M_{age} = 20.8 years) participated for course credit in this and an unrelated other experiment. All participants were Dutch speakers and had normal or corrected-to-normal vision. All participants gave informed consent before participation. The experiment took approximately 30 minutes to complete.

Materials. All stimuli were presented on a 19-inch CRT screen (75Hz, 1280 × 1024 pixels, white background color). A computer program written in Affect 4.0 (Spruyt, Clarysse, Vansteenwegen, Baeyens, & Hermans, 2010) was used to control stimulus presentation and response registration.

The reference objects used on the experimental trials were two black-and-white line drawings, one of a football (i.e. small reference object) and one a car (i.e. large reference object). The primes and the targets of the experimental trials were 36 black-and-white line drawings of objects of various sizes, sampled from a stimulus set developed by Severens, Lommel, Ratinickx, and Hartsuiker (2005). All stimuli had a resolution of 300 × 300 pixels (roughly 7.9 cm × 7.9 cm), corresponding with a visual angle of (about) 9.03 degrees. The stimuli belonged to one of three subsets based on their size relative to the two reference objects (see Figure 1). Twelve objects were smaller than both a car and a football (i.e., small stimuli; e.g., cup). Twelve objects were larger than both a football and a car (i.e., large stimuli; e.g., house). Finally, the twelve critical, medium-sized stimuli were larger than the football but smaller than the car (e.g., bike). For each individual participant, a prime set and a target set were sampled at random without replacement with the restriction that each stimulus set consisted of six small, six medium-sized, and six large stimuli. Stimuli that were used as targets were thus never used as primes (and vice versa) for an individual participant.[3]

For the practice trials, one reference object (i.e., a cow) and eight prime and target stimuli were used. Four stimuli depicted objects that were larger than a cow, whereas the other four stimuli depicted objects that were smaller than a cow. The stimuli used for the practice phase were also sampled from Severens et al. (2005). For each participant, both the subset of large and small objects were split at random into a prime set and a target set, each containing two small and two large stimuli each.

[3] The luminance of pictures showing small objects was more pronounced than the luminance of pictures showing large objects, $t(22) = 4.50$, $p < .001$, or medium-sized objects, $t(22) = 3.72$, $p < .005$. Given that stimuli from different stimulus categories were presented equally often in response-compatible and response-incompatible conditions, none of our critical effects can be attributed to variations in luminance.

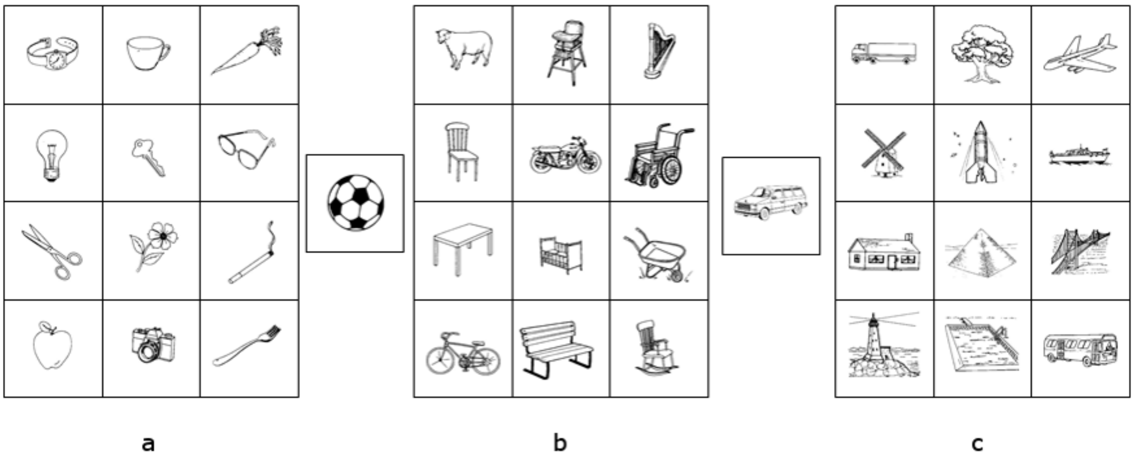

Figure 1. *Experimental stimuli, grouped in their respective size categories, and the two reference objects (i.e., the ball and the car). Small objects (a) are smaller than the two reference objects. Large objects (c) are larger than the two reference objects. Medium-sized objects (b) are smaller than the large reference object but larger than the small reference object*

Procedure. Participants were tested individually in a dimly-lit room. They were asked to categorize objects presented on the computer screen as larger than or smaller than a reference object. Participants were told that, on each trial, two objects would be presented in fast succession and were asked to categorize the second object. It was emphasized that the real-life sizes of the depicted objects should be the basis for their judgments. It was also mentioned that the reference stimulus would not change within blocks and that the meaning of the response keys would remain the same throughout the entire experiment (i.e., left key for targets smaller than the reference object, right key press for targets larger than the reference object). After an incorrect response, the sentence “Verkeerde reactie!” (Dutch for “Incorrect response!”) was presented for 1000 ms.

The practice phase comprised 16 trials. The experimental phase involved eight blocks of 60 trials each, leading to a total of 480 trials. The reference stimulus was the same throughout the first four blocks and was then changed for the last four blocks.

Participants were randomly assigned to one of two order conditions: They either first categorized target objects relative to a football and then to a car, or vice versa.

Trials can be divided into three general trial types: On 40 % of trials (i.e., 24 trials per block), both the primes and targets depicted large or small objects (hereafter referred to as *standard trials*). On 40 % of trials (i.e., 24 trials per block), the primes depicted medium-sized objects whereas the targets consisted of large and small objects (hereafter referred to as *critical trials*). On the remaining 20 % of trials (i.e., twelve trials per block), the primes depicted objects that were either small or large, whereas the targets consisted of medium-sized targets (hereafter referred to as *filler trials*). This third trial type was included because large and small targets were associated with the same response irrespective of the size of the reference object. By including trials with medium-sized targets, we ensured that participants were required to compare the targets relative to the relevant reference object. Within in each blocks, however, the same response was required towards all the medium-sized targets (i.e., always "larger than" or "smaller than"). Accordingly, to avoid an overall response bias within blocks, we adjusted the number of small and large targets within the standard trials. When participants categorized the target objects relative to a football (and medium-sized targets thus had to be judged as large on the twelve filler trials), small target objects were presented on 18 out of the 24 standard trials (per block). In contrast, when participants categorized the target objects relative to a car (and all medium-sized targets thus had to be judged as small on the twelve filler trials), large target objects were presented on 18 of the 24 standard trials (per block). Note, that this procedure had no effect on the congruency proportion within each block: On 50 % of the trials, the relation to the reference object of both the prime and the target was the same, whereas it was different on the other half of the trials. Trials were semi-randomly intermixed, controlling for the same percentage of standard, critical, and filler trials in each block. Presentations of the same target on consecutive trials were not allowed.

Each trial started with a 500 ms presentation of a fixation cross. Next, after an interstimulus interval of 500 ms, a prime picture was presented for 200 ms. Finally, 50 ms after the offset of the prime, a target picture was presented until a response was registered. The SOA was thus 250 ms. The next trial started after an interval that varied between 500 ms and 1500 ms.

Results

Data trimming. Practice trials and filler trials were excluded from the analyses. We also excluded all the data of two participants whose overall mean response latencies (758 ms and 767 ms) exceeded our cutoff criterion of 2.5 standard deviations above the grand mean (M = 511 ms, SD = 76 ms; threshold = 702 ms). Similarly, the data of two other participants were excluded because their error rates (14.8 % and 25.5 %) exceeded the cutoff criterion of 2.5 standard deviations above the grand mean (M = 3.1 %, SD = 3.8 %; threshold = 12.5 %).[4] At the trial level, we excluded all trials on which an incorrect response was registered (2.6 %).[5] Finally, the influence of outlier values was reduced by excluding all trials with response latencies that deviated more than 2.5 standard deviations from a participant's mean latency in a particular priming condition (3.0 %; Ratcliff, 1993).

Design. Two separate analyses were performed, one for the critical trials and one for the standard trials. The critical trials were subjected to a 2 (reference object: football vs. car) × 2 (target size: large vs. small) repeated measures ANOVA to check whether the influence of the medium-sized primes on the speed of responding was dependent upon the size of the reference object. Note, that the size of the reference object was used as a factor instead of the primes as the reference object was informative about the relative size of primes whereas the primes themselves were not. The RPE corresponds to the interaction between the size of the reference objects and the targets on critical trials. The standard trials were subjected to a 2 (*reference object*: football vs. car) × 2 (*prime size*: large vs. small) × 2 (*target size*: large vs. small) repeated measures ANOVA. It may be noted, however, that we expected automatic relational

[4] None of the critical effects in this experiment nor in the subsequently reported experiments were dependent on the exclusion or inclusion of participants.

[5] An analysis of the error data revealed effects that mirrored the effects found in the reaction time data. It must be noted, however, that the overall error rates were too small to allow for parametric testing (i.e., less than 2.7 %), especially when taking into account the complexity of the resign design. We therefore decided not to reported the error data in full detail.

stimulus processing to occur for standard and medium-sized primes alike. On standard trials, it is therefore impossible to disentangle the effects of automatic relational stimulus processing from the effects produced by the size relation between the primes and the targets in terms of their absolute sizes. Caution is thus in order when interpreting the magnitude of the priming effects observed on the standard trials.

Priming on critical trials. As expected, we obtained a significant interaction between the size of the reference objects and the targets, $F(1, 63) = 99.99$, $p < .001$, $\eta_\rho^2 = .61$. Large targets were categorized faster when the reference object was small than when the reference object was large (482 ms vs. 507 ms). Conversely, small targets were responded to faster when the reference object was large than when the reference object was small (477 ms vs. 498 ms). Overall, these results corresponded to a mean RPE of 23 ms. In addition, the main effect of target size was significant, $F(1, 63) = 4.83$, $p = .032$, $\eta_\rho^2 = .07$. Small targets were categorized faster than large targets (487 ms vs. 495 ms).

Priming on standard trials. As expected, the interaction between the size of the primes and the targets reached significance, $F(1, 63) = 35.35$, $p < .001$, $\eta_\rho^2 = .36$. Large targets were categorized faster when they were preceded by large primes than when they were preceded by small primes (487 ms vs. 503 ms). Conversely, responses towards small targets were faster after the presentation of small primes than after the presentation of large primes (488 ms vs. 502 ms). Overall, these results correspond with a priming effect of 15 ms. Three other effects were significant. First, the main effect of the size of the reference object was significant, $F(1, 63) = 9.09$, $p = .004$, $\eta_\rho^2 = .13$. Target categorizations were faster when the reference object was small compared to when the reference object was large (490 ms vs. 500 ms). Second, the size of the primes interacted significantly with the size of the reference object, $F(1, 63) = 4.11$, $p = .047$, $\eta_\rho^2 = .06$. When the reference object was large, the speed of responding was more or less the same after small and large primes (499 ms vs. 501 ms). Conversely, when the reference object was small, target responding was faster after large primes than after small primes (488 ms vs. 493 ms). Finally, the size of the targets interacted significantly with the size of the reference object, $F(1, 63) = 43.90$, $p < .001$, $\eta_\rho^2 = .41$. When the reference object was large, target responding was faster when the targets were small than when the targets were large (492 ms vs. 508 ms). Conversely, when the reference object was small, target responding was faster when the targets were large than when the targets were small (482 ms vs. 498 ms).[6]

Discussion

In line with our expectations, the RPE reached significance. This effect strongly suggests that participants processed the size of the primes relative to size of the reference object. Moreover, the present findings are consistent with the hypothesis that relational information processing can take place in an automatic fashion (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). First, given that participants were quite fast to respond (i.e., mean response latencies below 500 ms) and the SOA was short (i.e., 250 ms), it can be concluded that the RPE was produced by fast-acting processes. Second, given that the primes were task-irrelevant, it seems unlikely that the occurrence of the RPE is dependent upon the conscious intention to process the relative size of the prime objects. Finally, the present findings provide further support for the idea that task misapplication is an important determinant of the sequential priming effect. According to this viewpoint, prime processing takes place due to an unintentional misapplication of the target task to the prime stimuli. Accordingly, if the task requires relational judgments about the target, one can expect the primes to be processed in a similar fashion.

Somewhat unexpectedly, we also observed that target responding on the standard trials was dependent upon the relationship between the size of the reference objects and the size of the primes. More specifically, task performance was better when the

[6] As explained in the method section, the reference object was changed after four blocks and the order in which the two reference objects were used was counterbalanced across participants. We therefore examined whether target responding was affected by this counterbalancing factor. On critical trials there was evidence for a main effect of order, $F(1, 62) = 4.61$, $p = .036$, $\eta_p^2 = .07$. Overall, participants were faster to respond when the small reference object was used first compared to when the large reference object was used first (477 ms vs. 504 ms). Reassuringly, the RPE was not moderated by the order factor, $F < 1$, nor were any other effects on critical trials, all Fs < 1, all ps > .81. We therefore excluded the order factor from the analyses of Experiment 1. It may be noted, however, that there was some evidence that the counterbalancing factor did impact the prime × target interaction observed on the standard trials, $F(1, 62) = 3.52$, $p = .065$, $\eta_p^2 = .05$. The priming effect was larger when the small reference object was used first compared to when the large reference object was used first (20 ms vs. 11 ms). None of the other effects observed on the standard trials were qualified by the counterbalancing factor, all Fs < 1.40, all ps > .241.

size dissimilarity between the primes and the reference objects was large as compared to when it was small. The relationship between the reference objects and the primes thus seems to affect target responding in the same manner as the relationship between the reference objects and the targets. On the one hand, this similarity between the effects produced by the primes and the effects produced by the targets adds further weight to the idea that participants misapplied the target task to the prime set. On the other hand, it is unclear to us how a prime stimulus, once compared to the reference object, can affect target responding irrespective of the nature of the targets or the nature of the required response. It must be noted, however, that the interaction between the size of the reference objects and the size of the primes observed in Experiment 1 replicated only in Experiment 3. Moreover, a joint analysis of all 4 experiments revealed that the effect was not reliable across experiments, $F(1, 223) = 2.24$, $p = .14$, and was also not moderated by the experiment factor, $F(3, 223) = 1.76$, $p = .16$. We are therefore reluctant to put too much weight on this finding. Taking into account that this effect concerns the (non-critical) standard trials only, we will therefore not elaborate on this effect any further.

To summarize, our findings strongly suggest that relational information can be processed under automaticity conditions. Nevertheless, it is possible to identify two rivaling interpretations for our effects. First, it could be argued that the RPE observed on critical trials was an artifact resulting from the fact that the size of the reference object affected target responding in a direct, non-relational manner. Second it might be argued that our effects reflected memory retrieval processes rather than from automatic relational information processing. We now discuss these rivaling interpretations one by one.

Inflation of the RPE. As explained above, we expected the evaluation of the size of the medium-sized primes to depend on the size of the reference object. More specifically, medium-sized primes were expected to be evaluated automatically as being smaller than the large reference object and larger than the small reference object. We therefore used the interaction between the size of the reference objects and the targets on critical trials as an index of automatic relational processing. One might object, however, that target responding was affected directly by the relationship between the size of the references object and the targets, that is, irrespective of the primes. According to this viewpoint, target responding was simply faster for large targets when the reference object was small compared to when it was large, and vice versa for small targets. To account for such an effect, it could be argued that it is simply easier to determine the relative size of an object when the difference between that object and the reference object is large (i.e., the *distance effect*; Banks, 1977; Schneider & Logan, 2007).

It is important to point out, however, that there are no reasons to suspect that the size of such an effect would be different on standard trials as compared to critical trials. Therefore, a comparison of the interaction of the size of the reference objects and the targets on both types of trials can shed light on the extent to which the RPE was truly due to relational processing. We thus performed an additional analysis in which all critical and standard trials were subjected to a 2 (*trial type*: standard vs. critical) × 2 (*reference object*: football vs. car) × 2 (*target size*: large vs. small) repeated measures ANOVA. Reassuringly, the three-way interaction between trial type, reference object, and target size reached significance, $F(1, 63) = 4.57$, $p = .036$, $\eta_p^2 = .07$, showing that the interaction between the size of reference objects and the targets was significantly smaller on standard trials (i.e., 16 ms) as compared to critical trials (i.e., 23 ms, the RPE). Accordingly, it can be concluded that we obtained clear evidence for the occurrence of automatic relational stimulus processing, even though target responding was also influenced by the relationship between the target and the reference object. To corroborate this interpretation even further, we also decided to add an additional control condition to the design of Experiment 2 (see below).

Automatic relational processing vs. automatic memory retrieval. As argued above, there are good reasons to argue that the RPE observed in the present study resulted from automatic relational stimulus processing. One could also argue, however, that automatic memory retrieval was the driving force behind our effects. Remember that the size of the reference object was manipulated on a block-wise basis. That is, participants first completed four blocks of trials with one reference object and then completed four additional blocks of trials with the second reference object. Under these circumstances, it seems difficult to rule out the possibility that participants processed the primes in a non-automatic fashion on at least a number of trials and then retrieved the outcome of this prime-evaluation process on subsequent trials. Such an interpretation implies, however, that the magnitude of the RPE should have increased across blocks. Reassuringly, additional analyses confirmed that the interaction between reference object and target size did not differ between blocks, $F < 1$. Nevertheless, we decided to perform a second study in which this potential problem was addressed empirically rather than statistically.

EXPERIMENT 2

Experiment 2 was a conceptual replication of Experiment 1, aimed at ruling out two rivaling interpretations of our findings. First, to rule out the possibility that the RPE was driven by (automatic) memory retrieval rather than automatic relational stimulus processing, we manipulated the size of the reference object on a trial-by-trial basis. As a result, even if participants retrieved the outcome of the prime-evaluation process on preceding trials, automatic memory retrieval could no longer exert a systematic influence on the RPE. Second, we included a number of trials on which scrambled versions of the medium-sized objects were used as primes. As explained above, one might argue that target responding slows down as the absolute size of the targets and the reference objects becomes more similar. Because such a distance effect could inflate the RPE, it was key to demonstrate that the magnitude of the reference object × target interaction observed on trials with unscrambled medium-sized primes (i.e., the RPE + the distance effect) was substantially larger as compared to trials with scrambled medium-sized primes (i.e., the distance effect).

Method

Participants. Fifty-five students at Ghent University (42 women, 8 men; M_{age} = 21.45 years) participated in this 30-minute experiment and received 5€ payment. All participants were Dutch speakers and had normal or corrected-to-normal vision. All participants gave informed consent before participation.

Materials. All materials were identical to those used in Experiment 1, with two exceptions. First, next to the original set of medium-sized images, we now included scrambled versions of these stimuli. Second, stimuli were presented on a 20-inch LCD screen (75Hz, 1680 × 1050 pixels) instead of 19-inch CRT screen. As a result, the physical size of the pictures was roughly 8.5 cm by 8.5 cm, corresponding with a visual angle of (about) 9.72 degrees.

Procedure. The procedures used in the present experiment were identical to those used Experiment 1, with four exceptions. First, the size of the reference object was manipulated on a trial-wise basis. Each trial started with a 1000 ms presentation of one of the reference stimuli as a cue. Next, after an interstimulus interval of 500 ms, the prime picture was presented for 200 ms. Finally, 50 ms after the offset of the prime (SOA 250 ms), the target picture was presented at the same location until participants responded by pressing one of the two response keys. Participants were told that the reference stimuli would change from trial to trial. They were asked to compare the size of the targets relative to the preceding reference object. Second, the line drawings of a car and a football were now used as the reference stimuli both on the practice trials and the experimental trials. Consequentially, three of the eight practice stimuli were dropped, and eleven other stimuli were added, leading to six small stimuli, six large stimuli, and four medium-sized stimuli. Third, we included trials in the experimental phase on which the primes were scrambled versions of the medium-sized objects and the targets depicted large or small objects (hereafter referred to as *neutral trials*). Fourth, the number of trials was reduced to 360 trials (four blocks of 90 trials each). Each block consisted of 24 standard, 24 critical, 24 neutral, and 18 filler trials. Small and large targets were counterbalanced across trial types. Medium-sized targets were presented four times within each block of trials: twice on a trial with the small reference object and twice on a trial with a large reference object. Within each block of trials, the relatedness proportion was 50 %. Trials were semi-randomly intermixed, controlling for the above mentioned distribution of standard, critical, neutral and filler trials within blocks.

Results

Data trimming. Due to a computer malfunction, the data set of one participant was not complete and was therefore excluded from the analyses. The criteria for excluding participants and trials were identical to those in Experiment 1. Specifically, we excluded the data of two participants whose overall mean response latencies (762 ms and 771 ms) exceeded our cutoff criterion of 2.5 standard deviations above the grand mean (*M* = 548 ms, *SD* = 84 ms; threshold = 757). The data of three other participants were excluded because their error rates (7.6 %, 8.7 % and 8.7 %) exceeded the cutoff criterion of 2.5 standard deviations above the grand mean (*M* = 2 %, *SD* = 2 %; threshold = 7 %). At the trial level, we again excluded all trials on which an incorrect response was registered (1.7 %). Finally, to reduce the influence of outlier values, we excluded all trials on which a response latency was observed that deviated more than 2.5 standard deviations from a participant's mean latency in a particular priming condition (2.9 %; Ratcliff, 1993).

Design. To verify whether the interaction between the size of the reference objects and the targets was significantly larger on critical trials as compared to neutral trials, critical and neutral trials were subjected to a 2 (*trial type*: critical vs. neutral) × 2 (*reference object*: football vs. car) × 2 (*target size*: large vs. small) repeated measures ANOVA. The standard trials were analyzed in the same way as in Experiment 1.

Priming on critical and neutral trials. A strong interaction between the size of the reference objects and the size of the targets was found, $F(1, 48) = 27.75$, $p < .001$, η_ρ^2 = 37. Large targets were responded to faster when the reference object was small than when the reference object was large (519 ms vs. 530 ms). Conversely, small targets were responded to faster when the reference object was large than when the reference object was small (512 ms vs. 527 ms). Overall, these results corresponded to an interaction effect of 13 ms. Crucially, this 2-way interaction was further qualified by a three-way interaction between trial type, reference object, and target, $F(1, 48) = 6.14$, $p = .017$, $\eta_\rho^2 = .11$. We therefore performed separate analyses for the neutral trials and the critical trials.

On the subset of critical trials, the interaction between the size of the reference objects and the size of the targets was highly reliable $F(1, 48) = 32.14$, $p < .001$, $\eta_p^2 = .40$. Large targets were categorized faster when the reference object was small than when the reference object was large (519 ms vs. 533 ms). Conversely, small targets were responded to faster when the reference object was large than when the reference object was small (508 ms vs. 531 ms). Overall, these results corresponded to a mean RPE of 19 ms.

On the subset of neutral trials, the interaction between target and reference object just missed conventional significance levels, $F(1, 48) = 3.68$, $p < .061$, $\eta_p^2 = .07$. Large targets were categorized faster when the reference object was small than when the reference object was large (519 ms vs. 527 ms). Conversely, small targets were responded to faster when the reference object was large than when the reference object was small (517 ms vs. 522 ms). Overall, the effect of the size relation between the reference objects and the targets amounted to 7 ms.

Priming on standard trials. As expected, the interaction between prime size and target size reached significance, $F(1, 48) = 10.97$, $p = .002$, $\eta_p^2 = .19$. Large targets were categorized faster when they were preceded by large primes than when they were preceded by small primes (521 ms vs. 534 ms). Conversely, responses towards small targets were faster after the presentation of small primes than after the presentation of large primes (519 ms vs. 533 ms). Overall, these results corresponded to a priming effect of 13 ms. The reference object × target interaction on standard trials was not significant, $F < 1$, nor descriptively different from zero. Moreover, the reference object × target interaction on standard trials was significantly smaller than the RPE (0 ms vs. 19 ms), $F(1, 48) = 18.68$, $p < .001$, $\eta_p^2 = .28$.

Discussion

The results are clear-cut. Replicating the findings of Experiment 1, a reliable RPE emerged. Crucially, the degree to which the size relation between the reference objects and the targets exerted an influence on target responding was significantly larger on the critical trials (i.e., 19 ms) as compared to the neutral trials (i.e., 7 ms). Given that the critical trials and the neutral trials were identical in terms of the reference objects and the target stimuli that were used, this data pattern could be due only to the fact that the relational processing of the primes exerted an influence on target responding on the critical trials. We can therefore safely conclude that the RPE is not just a byproduct of the fact that it is easier to judge the relative size of an object as the difference in size between the reference object and the target object becomes larger.[7]

The results of Experiment 2 are also important for a second reason. As pointed out earlier, memory retrieval may have contributed to the emergence of the RPE in Experiment 1. According to this viewpoint, participants processed the relative size of the primes in a non-automatic fashion on at least a number of trials and then retrieved the outcome of this deliberate prime-evaluation process on subsequent trials. In the present experiment, however, such an account is less likely because the size of the reference objects was now manipulated on a trial-wise basis instead of a block-wise basis. Even if knowledge about the size relation between the medium-sized primes and the reference objects was stored in episodic memory, the trial-wise variation of the reference objects renders it rather unlikely that the (automatic) retrieval of this information could influence target responding in a systematic way. Nevertheless, one might argue that a constant change of the reference object does not completely exclude the possibility that memory retrieval processes contributed to the emergence of the RPE. Therefore, additional analyses were performed. If participants established S-R bindings first and then activated them on subsequent trials, one would predict the RPE to increase as a function of the number of prime repetitions. Reassuringly, these analyses revealed no supporting evidence for this prediction, $F < 1$.

Finally, one might note that the distance effect observed on the standard trials (i.e., the interaction between the size of the reference objects and the size of the targets) was absent in the present experiment whereas the same interaction did reach

[7] Experiment 2 was a replication of an earlier study, the results of which mimic those reported here. Specifically, we obtained a significant RPE of 16 ms, $F(1, 56) = 12.69$, $p = .001$, $\eta_p^2 = .18$, whereas the reference object × target interaction on the standard trials was non-significant, $F(1, 56) = 1.73$, $p = .194$, $\eta_p^2 = .03$. Crucially, we did not include neutral trials in this study. Therefore, the RPE was compared to the reference object × target interaction on standard trials to check for a possible inflation of the RPE. The three-way interaction between the size of the reference objects, the size of the targets, and the trial type failed to reach significance, $F(1, 56) = 1.70$, $p = .197$, $\eta_p^2 = .03$.

significance in Experiment 1. Additional analyses confirmed that this difference in the magnitude of the distance effect was statistically reliable, $F(1, 111) = 13.93$, $p < .01$. In all likelihood, this data pattern resulted from the fact that the reference objects were manipulated block-wise in Experiment 1 but varied randomly from trial to trial in Experiment 2. Irrespectively, the fact that the distance effect was absent in the present experiment whereas the RPE was not only adds to the idea that the RPE is not just a byproduct of the distance effect.[8]

EXPERIMENT 3

Experiments 1 and 2 provided convincing evidence for the hypothesis that relational stimulus information about the relative size of an object can be processed in an automatic fashion. Experiment 3 was primarily designed to test the hypothesis that the RPE arises as the result of a goal-dependent but unintentional misapplication of the target task to the prime set (Dehaene et al., 1998; Van Opstal et al., 2010). In theory, the likelihood of task-misapplication should decrease under conditions that promote a clear separation of the primes and the targets. Accordingly, we manipulated whether the (critical) medium-sized primes were presented in the same or a in a different color as the target stimuli and expected the RPE to be less pronounced in the color-incongruent condition relative to the color-congruent condition.

Experiment 3 was also important for an additional reason. Whereas the color of the medium-sized primes was manipulated on a within-subject basis, the color of the targets was manipulated on a between-subjects basis. We deliberately opted for this strategy to ensure that the critical priming trials were all identical for a given participant, except for the presence or absence of a color match/mismatch of the primes. Accordingly, should the magnitude of the RPE vary as a function of stimulus similarity, one can safely conclude that the RPE is not just a byproduct of the distance effect described earlier.

[8] It may also be noted that, in Experiment 1, the standard trials revealed a reliable main effect of the size of the reference objects whereas the same effect was unreliable in Experiment 2. Additional analyses provided no evidence for the hypothesis that the magnitude of this effect was reliably different in both experiments, $F(1, 111) = 2.02$, $p = .16$.

Method

Participants. Seventy-one students at Ghent University (54 women, 17 men; M_{age} = 21.06 years) participated in this 30-minute experiment and received 5€ payment. All participants were Dutch speakers and had normal or corrected-to-normal vision. All participants gave informed consent before participation.

Materials. The stimulus materials used in the present experiment were identical to those used in Experiment 1. However, all primes and targets were now presented in orange or blue. Similar to Experiment 2, the experiment was conducted using a 20-inch LCD screen (75Hz, 1680 × 1050 pixels). Accordingly, the physical size of the experimental stimuli was roughly 8.5 cm by 8.5 cm, corresponding with a visual angle of (about) 9.72 degrees.

Procedure. The procedure was identical to that used in Experiment 1, with two exceptions. First, the color of the medium-sized primes was manipulated trial-by-trial on a within-subjects basis (i.e., either orange or blue). Second, the color of all the other stimuli was manipulated on a between-subjects basis (i.e., either orange or blue). Crucially, on 50 % the critical trials, the primes and the targets were presented in the same color. On the remaining trials, the primes and the targets were presented in a different color. Given that critical trials represented on 40 % of all trials (see Experiment 1), the color of the prime and the target differed on 20 % of all trials.

Results

Data trimming. The criteria for excluding participants and trials were identical to those in Experiment 1. Specifically, we excluded the data of one participant whose overall mean response latencies (964 ms) exceeded our cutoff criterion of 2.5 standard deviations above the grand mean (*M* = 517 ms, *SD* = 88 ms; threshold = 738 ms). We also excluded the data of one participant whose error rate (12.2 %) exceeded the cutoff criterion of 2.5 standard deviations above the grand mean (*M* = 3.1 %, *SD* = 3.3 %; threshold = 11.2 %). Finally, we excluded the data of a third participant who was an extreme outlier both in terms of mean response latency (784 ms) as well as overall error rate (18 %). At the trial level, we again excluded all trials on which an incorrect response was registered (2.7 %). Finally, to reduce the influence of outlier values, we excluded all trials on which a response latency was observed that deviated more than 2.5 standard deviations from a participant's mean latency in a particular priming condition (3.0 %; Ratcliff, 1993).

Design. Critical and standard trials were analyzed as in Experiment 1, with the exception that for the analysis of critical trials the factor color congruency was added to the design to check for the influence of color congruency of prime and target on the RPE.[9]

Priming on critical trials. We found a main effect of target size, $F(1, 66) = 11.27$, $p = .001$, $\eta_p^2 = .15$, which was further qualified by size of the reference object, $F(1, 66) = 85.21$, $p < .001$, $\eta_p^2 = .56$. Large targets were categorized faster when the reference object was small than when the reference object was large (480 ms vs. 516 ms). Conversely, small targets were responded to faster when the reference object was large than when the reference object was small (477 ms vs. 497 ms). Overall, these results corresponded to a mean RPE of 28 ms. Crucially, color congruency qualified the interaction between the size of the reference objects and the targets, $F(1, 66) = 19.66$, $p < .001$, $\eta_p^2 = .23$. In contrast to our expectations, however, the mean RPE was larger when the color of the primes and the targets did not match than when the primes and the targets were presented in the same color (32 ms vs. 24 ms). This three-way interaction was further qualified by the order of the reference objects, $F(1, 66) = 8.91$, $p = .004$, $\eta_p^2 = .12$. When participants first categorized targets relative to the small reference object, the difference between the mean RPE observed on color incongruent

[9] As in Experiment 1, the reference object was changed after four blocks. The order in which the two reference objects were used as well as the color in which the majority of the stimuli were presented were counterbalanced across participants. None of the critical effects were moderated by the color of the stimuli on neither the critical nor the standard trials, all *F*s < 1.43, all *p*s > .236. The color factor was thus excluded from further analyses. However, the order in which the reference objects were used did affect critical effects. The influence of the order of the reference object on the RPE just missed conventional significance levels, $F(1, 66) = 3.96$, $p < .051$, $\eta\rho 2 = .06$. When participants first categorized targets relative to the small reference object, the mean RPE was 33 ms. Conversely, when participants were first presented with the large reference object, the mean RPE was 23 ms. Moreover, the order factor qualified the critical three-way interaction between color congruency, size of the reference object, and size of the targets on critical trials, $F(1, 66) = 8.91$, $p = .004$, $\eta\rho 2 = .12$. The order factor was thus kept for the analysis of the critical trials.

trials and color congruent trials was larger (41 ms vs. 26 ms) than when participants were first presented with the large reference object (24 ms vs. 23 ms).

Priming on standard trials. As expected, the interaction between the size of the primes and the targets reached significance, $F(1, 67) = 46.70$, $p < .001$, $\eta_\rho^2 = .41$. Large targets were categorized faster when they were preceded by large primes than when they were preceded by small primes (488 ms vs. 501 ms). Conversely, responses towards small targets were faster after the presentation of small primes than after the presentation of large primes (483 ms vs. 502 ms). Overall, these results corresponded to a priming effect of 16 ms. Three other effects were significant. First, the main effect of the size of the reference object was significant, $F(1, 67) = 6.44$, $p = .014$, $\eta_\rho^2 = .09$. Target categorizations were faster when the reference object was small compared to when the reference object was large (489 ms vs. 499 ms). Second, the size of the prime interacted significantly with the reference object, $F(1, 67) = 12.84$, $p = .001$, $\eta_\rho^2 = .16$. When the reference object was large, target responding was faster after a small prime than after a large prime (494 ms vs. 503 ms). Conversely, when the reference object was small, target responding was faster after a large prime than after a small prime (488 ms vs. 490 ms). Finally, the size of the targets interacted significantly with the size of the reference object, $F(1, 67) = 60.89$, $p < .001$, $\eta_\rho^2 = .48$. When the reference object was large, target responding was faster when the targets were small than when they were large (488 ms vs. 509 ms). Conversely, when the reference object was small, target responding was faster when the targets were large than when the targets were small (480 ms vs. 497 ms). These results corresponded to an effect of 19 ms. However, the interaction was significantly smaller than the RPE of 28 ms, $F(1, 67) = 8.86$, $p = .004$, $\eta_\rho^2 = .12$.

Discussion

As in our previous experiments, the critical trials revealed a significant interaction between the size of the reference objects and the targets. Moreover, the magnitude of this RPE was larger on color-incongruent trials as compared to color-congruent trials. This finding is crucial as it adds further weight to the idea that the RPE is driven by the automatic relational processing of the primes. With the exception of the color of the primes, color-congruent trials and color-incongruent trials were identical for a given participant. Therefore, the observation that the RPE was larger on color-incongruent trials as compared to color-incongruent trials can only be attributed to differences in the extent to which relational information about the primes was processed.

It must be noted, however, that the nature of this interaction effect was unexpected. As explained above, we expected the likelihood of task misapplication (and hence the likelihood of obtaining a RPE) to be larger on color-congruent trials as compared to color-incongruent trials. The observed data pattern was exactly the opposite, but still it can be readily accounted for. Remember that all stimuli except the medium-sized primes were presented in the same color throughout the entire experiment for a given participant. In contrast, the color of the medium-sized primes varied randomly from trial to trial. As a result, the color used for the presentation of the color-incongruent primes was a rare event. It is well-known that infrequent events can attract attention because they are more salient than events that occur frequently (e.g., Barcelo, Escera, Corral, & Periáñez, 2006; Downar, Crawley, Mikulis, & Davis, 2000; Parkhurst, Law, & Niebur, 2002). It can thus be hypothesized that the color-incongruent primes were processed more deeply than the color-congruent primes because they were more salient and captured more attention. Additional research would be required, though, to firmly substantiate the hypothesis that involuntary task misapplication is dependent upon stimulus salience.

Finally, it may be noted that the distance effect (i.e., the interaction between the size of the reference objects and the size of the targets observed on the standard trials) as well as the main effect produced by the reference objects on the standard trials reached significance. Remember that both effects were also observed in Experiment 1 but were absent in Experiment 2. As argued earlier, this data pattern is likely to have resulted from the fact that the reference objects were manipulated block-wise in Experiments 1 and 3 but varied randomly from trial to trial in Experiment 2. Importantly, both effects are by definition unrelated to the RPE observed on the critical trails. Accordingly, we will not elaborate upon them any further.

EXPERIMENT 4

As discussed above, the studies by Van Opstal et al. (2010) revealed that participants can detect the presence or absence of an equivalence relation between two stimuli even if the instigating stimuli are presented below subjective recognition thresholds. It remains an open question, however, whether the analysis of more complex relational information is contingent upon stimulus awareness. In Experiment 4, we therefore sought to replicate the RPE using prime stimuli that were presented below subjective recognition thresholds. In addition, the observation of a subliminal RPE would strengthen our position that relational prime processing is truly unintentional.

To ensure that the primes were presented below subjective recognition thresholds, the prime duration used in Experiment 4 was reduced to just 23.5 ms. In addition, we used both a forward and a backward mask. Finally, to test whether our masking procedure was successful, we implemented a direct prime discrimination task. The regression method developed by Draine and Greenwald (1998) was then used to test for the significance of the RPE in the absence of prime awareness.

Method

Participants. A total of 50 students at Ghent University (40 women, 10 men; M_{age} = 20.98 years) participated for course credit or were paid 8€ for their help in this and another experiment. Each experiment took approximately 30 minutes. Twenty-nine students executed this experiment first, 21 started with an unrelated other experiment. All participants were Dutch speakers and had normal or corrected-to-normal vision. All participants gave informed consent before participation.

Materials. All stimulus materials were the same as those used in Experiment 2, with the exception that scrambled versions of the medium-sized stimuli were not used. The experiment was conducted using a 19-inch CRT screen (85Hz, 1280 × 1024 pixels). Accordingly, as in Experiment 1, the physical size of the pictures was (approximately) 7.9 cm by 7.9 cm, corresponding with a visual angle of (about) 9.03 degrees.

Procedure. The procedure used in the present experiment was identical to that used in Experiment 2, with the following three exceptions. First, to reduce the number of trials, the neutral trial condition was omitted from the design. The priming phase of experiment included 240 trials in total (i.e., four blocks of 24 critical trials, 24 standard trials, and twelve filler trials each). Second, the sequence of events on each trial was adjusted to ensure that the primes were presented subliminally. Each trial started with a 1000-ms presentation of a reference stimuli. Next, after an interstimulus interval of 300 ms, a forward mask (i.e., a random distribution of black, grey, and white pixels with a resolution of 300 × 300 pixels) was presented for 300 ms. A prime picture was then presented for 23.5 ms, immediately followed by a 35-ms presentation of a backward mask (identical to the forward mask). Finally, at the offset of the backward mask (i.e., SOA = 58.5 ms), the target picture was presented at the same location as the primes until participants responded by pressing one of the response keys. As a third procedural change relative to Experiment 2, participants completed a direct prime-recognition task after completion of the priming task. This prime-recognition task was identical to the main priming task, with the exception that participants were now asked to categorize the size of the primes relative to the size of the reference object. They were asked to respond as accurately as possible and were informed that it was no longer necessary to respond as fast as possible. Participants were asked to guess if they were unable to identify a prime stimulus. No error feedback was given.

Results

Data trimming. The criteria for excluding participants and trials were identical to those in Experiment 1. Specifically, we excluded the data of three participants whose overall mean response latencies (968 ms, 1034 ms, 1187 ms) exceeded our cutoff criterion of 2.5 standard deviations above the grand mean (M = 603 ms, SD = 141 ms; threshold = 955 ms). We also excluded the data of one other participant whose error rate (5.7 %) exceeded the cutoff criterion of 2.5 standard deviations above the grand mean (M = 1.1 %, SD = 1.2 %; threshold = 4.2 %). At the trial level, we again excluded all trials on which an incorrect response was registered (1.0 %). Finally, to reduce the influence of outlier values, we excluded all trials on which a response latency was observed that deviated more than 2.5 standard deviations from a participant's mean latency in a particular priming condition (2.8 %; Ratcliff, 1993).

Design. The response latency data were subjected to the same set of analyses that were performed for Experiment 2, with one exception. Because the neutral prime condition was removed from the study design, the magnitude of the RPE was assessed relative to magnitude of the interaction between the size of the reference objects and the size of the targets observed on standard trials (as in Experiments 1 and 3).

Priming on critical trials. The interaction between the size of the reference objects and the size of the targets again reached significance, $F(1, 45) = 6.71$, $p = .013$, $\eta_p^2 = .13$. Large targets were categorized faster when the reference object was small than when the reference object was large (556 ms vs. 567 ms). Conversely, small targets were responded to faster when the reference object was large than when the reference object was small (541 ms vs. 561 ms). Overall, these results corresponded with a mean RPE of 10 ms. In addition, a significant main effect of target size was found, $F(1, 45) = 5.15$, $p = .028$, $\eta_p^2 = .10$. Small targets were categorized faster than large targets (549 ms vs. 564 ms).

Priming on standard trials. No significant effects were found, all Fs < 3.58, all ps > .065. Most importantly, both the interaction between prime size and target size and the interaction between the size of the reference objects and the size of the targets failed to reach significance, $F < 1$ and $F(1, 45) = 1.03$, $p = .316$, $\eta_p^2 = .02$, respectively. It might be noted, though, that both effects were numerically in the expected direction (i.e., 3 ms and 5 ms, respectively). Importantly, the strength of the interaction between the size of the reference objects and the size of the targets was roughly the same on critical priming trials (i.e., the RPE, 10 ms) and standard priming trials (i.e., 5 ms), $F(1, 45) = 1.11$, $p = .297$, $\eta_p^2 = .02$.

Regression analysis. Linear regression analyses were performed to establish the subliminal character of the RPE (Draine & Greenwald, 1998). To that end, we first calculated individual *d'* scores on the basis of the prime discrimination task. The proportion of correct prime categorizations was used as the hit rate whereas the proportion of incorrect prime categorizations was used as the false alarm rate (Green & Swets, 1966). On average, prime perceptibility was significantly different from zero, M = 0.34, $t(44) = 8.58$, $p < .001$, $d = 1.81$. The individual RPE scores were then regressed on the individual *d'* scores. The intercept in such a regression analysis indicates whether the RPE occurred in the absence of prime awareness. The slope, on the other hand, captures the relation between prime visibility and the RPE. One participant was identified as an influential case(i.e., Cook's distance p > .20; Kutner, Nachtsheim, Neter, & Li, 2005) and was therefore excluded from the regression analyses. It may be noted that the results reported below would improve dramatically if this case were not excluded. It is thus safe to say that we opted for a conservative approach.

Figure 2 depicts the linear regression line and a scatterplot of the individual data points. Both the intercept (i.e., 7.17 ms) and the slope (i.e., 11.19 ms) were non-significant, *t*s < 1.37, *p*s > .15. Virtually identical findings emerged when using the improved regression approach that was proposed by Klauer, Draine, and Greenwald (1998) to accommodate measurement error in the prime visibility regression predictor (see also Klauer, Greenwald, & Draine, 1998).

Discussion

While the overall RPE did reach significance, the present experiment failed to produce convincing evidence for the hypothesis that automatic relational processing can take place even if participants are completely unaware of the instigating stimulus. First, the mean *d'* score was statistically different than zero, suggesting that our masking procedure may have been ineffective for at least a subset of participants. Second, a regression analysis in which the individual RPE scores were regressed on the individual *d'* scores revealed non-significant results only. Third, the strength of the interaction between the size of the reference objects and the size of the targets was more or less the same on critical priming trials (i.e., the RPE) and standard priming trials. In sum, the results of Experiment 4 are insufficient to substantiate the hypothesis that automatic relational stimulus processing can take place in the absence of stimulus awareness. It is possible that the specific experimental setup (i.e., two reference objects, a large and unpredictable prime stimulus set, short SOA) prevented the occurrence priming effects (e.g., Damian, 2001; Kunde et al., 2003). On the other hand, it may be noted, that (a) the all effects were numerically in the anticipated direction and (b) the sample size of the present experiment was quite limited. It might thus be worthwhile to replicate the present experiment while ensuring adequate statistical power.

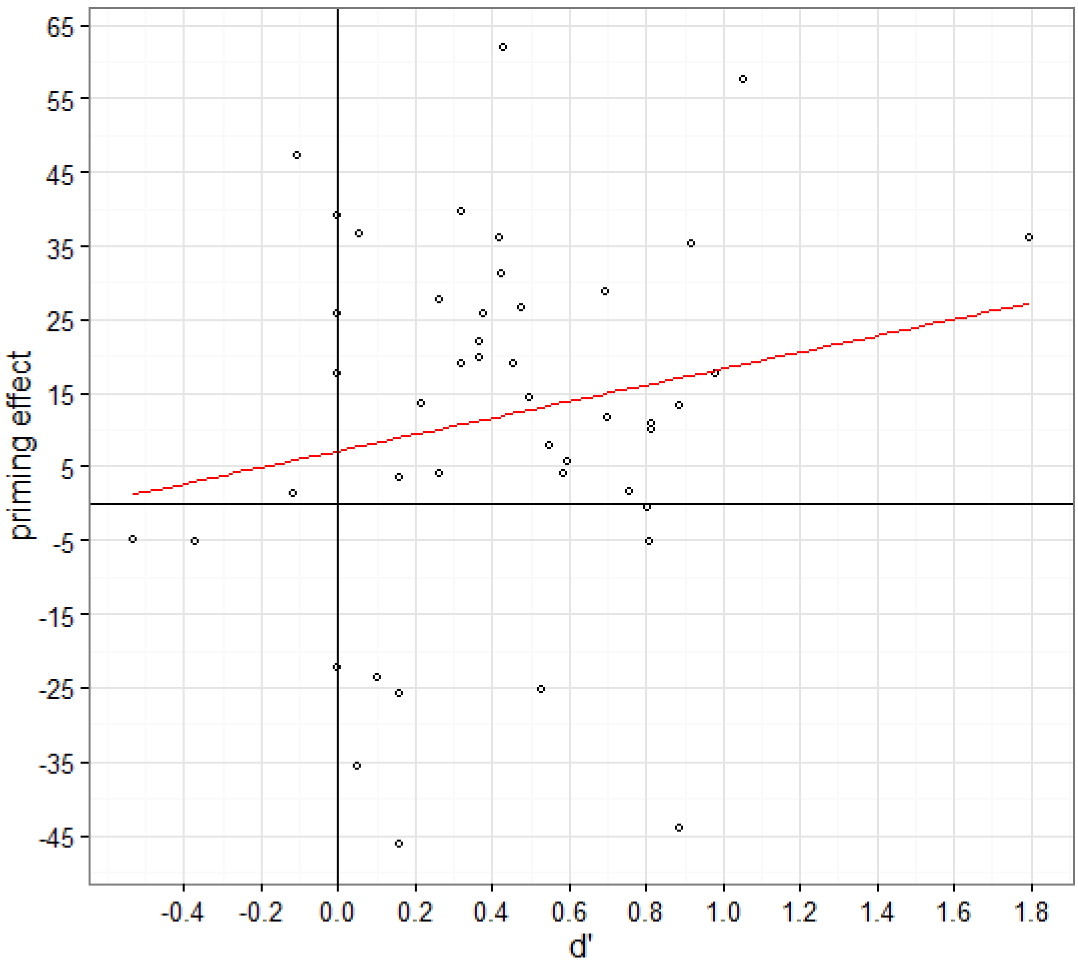

Figure 2. *Regression of the relational priming effect on individual d' measures of prime perceptibility: Scatterplot and best fitting regression function*

GENERAL DISCUSSION

To examine whether complex relational information can be processed under automaticity conditions, we conducted four sequential priming studies in which participants were asked to judge the physical size of a series of target objects relative to either a small reference object (i.e., a football) or a large reference object (i.e., a car). Crucially, the prime set included objects that were larger than the small reference object and smaller than the large reference object (e.g., a bike). Results showed that the impact of these medium-sized objects on target categorization was contingent upon the size of the reference object. That is, we observed that the medium-sized primes facilitated responses towards large targets and inhibited responses towards small targets when the size of the reference object was small. Conversely, when the size of the reference object was large, the same set of medium-sized primes facilitated responses towards small targets and inhibited responses towards large targets. This RPE was obtained when the size of the reference object was manipulated block-wise (Experiments 1 and 3) and trial-wise (Experiments 2 and 4). We also found some evidence that the RPE depends on the salience of the prime stimuli (Experiment 3). Finally, results were inconclusive as to whether the RPE might replicate in the absence of prime awareness (Experiment 4). The implications of these findings are twofold. First, our findings corroborate the hypothesis that complex relational information can be processed under automaticity conditions. Second, our findings shed new light on the mechanisms that are at play in the sequential priming paradigm. We will discuss these implications one by one.

Complex relational stimulus processing under automaticity conditions?

The mere fact that an RPE was found in each of our experiments is sufficient to conclude that participants were able to process complex relational prime information. The question to be addressed here concerns the extent to which the conditions implemented in our studies allow for the conclusion that this type of prime processing can take place under automaticity conditions. Reassuringly, a careful analysis of our experimental procedures in terms of the automaticity features put forward by Moors and De Houwer (2006) strongly supports the conclusion that the RPE is indeed produced by automatic processes. First, it seems highly unlikely that the RPE is dependent upon the (conscious) intention to process the size relation between the primes and the reference objects. Not only were participants asked to respond on the basis of the targets only, the prime information was also uninformative for the task at hand. Moreover, in Experiment 3, it was observed that the RPE is more pronounced for salient than for non-salient primes. It seems highly unlikely that participants would limit their intentional attempts to extract task-relevant prime information to just one subset of experimental stimuli. The findings of Experiment 3 therefore add further weight to the idea that automatic relational stimulus processing can occur unintentionally. Second, both the SOA and the presentation time of the primes were quite short in each of our experiments. In addition, participants were always very fast to respond (i.e., 495 ms, 524 ms, 496 ms, and 559 ms in Experiments 1, 2, 3, and 4 respectively). It can thus be concluded that the RPE observed in each of our experiments was driven by fast-acting processes. Third, the response task implemented in the present series of experiments was much more demanding than the response tasks typically used in sequential priming research. It could thus be argued that the RPE emerged under conditions of reduced cognitive capacity, implying that the processes driving the RPE must be efficient. On the other hand, the results of Experiment 4 regarding the need for awareness of the prime stimuli were inconclusive. Although an RPE was observed in Experiment 4, we were unable to rule out the possibility that the magnitude of the RPE was inflated by the distance effect. Moreover, while the prime duration was reduced to just 23.5 ms, a regression analysis failed to produce convincing evidence for the idea that the RPE can replicate even if participants are completely unaware of the primes. Further research would thus be needed to firmly substantiate the hypothesis that the RPE is not dependent upon the conscious identification of the primes. Taken together, however, our findings provide strong corroborating evidence for the hypothesis that complex relational stimulus processing can take place under at least three automaticity conditions (see Moors & De Houwer, 2006).

The mechanism underlying the involuntary processing of task-irrelevant stimuli

In line with Dehaene et al. (1998) and Van Opstal et al. (2010), we assume that automatic relational stimulus processing in the sequential priming paradigm results from a goal-dependent but unintentional application of the target task to the prime stimuli. Earlier research already provided supporting evidence for the hypothesis that the automatic analysis of task-irrelevant stimuli is confined to stimulus features that are task-relevant. Kiefer and Martens (2010), for example, demonstrated that the semantic analysis of briefly presented prime words is prevented if a non-semantic task set is activated prior to the presentation of the primes (for related findings, see Ansorge & Heumann, 2003; Ansorge et al., 2002, 2010, 2009, 2014; Ansorge & Neumann, 2005; Kiefer & Martens, 2010). The present research extends these findings by showing that task-misapplication (or “task-set execution” in the terminology used by Ansorge et al., 2014) can result not only in a top-down selection of specific features of a single stimulus, but also in a comparative integration of various sources of information.

Importantly, the explanatory power of a task misapplication account is by no means restricted to priming studies aimed at capturing automatic relational processing (e.g., Van Opstal et al., 2010). It can readily account for any type of priming effect whenever there is dimensional overlap between the prime set and the target set (Kornblum, Hasbroucq, & Osman, 1990). For instance, if the task requires participants to process the evaluative connotation of the targets, participants are likely to process the evaluative tone of the primes as well. Accordingly, one can expect the evaluative match between the primes and the targets to affect target responding in such a task, as was observed in an overwhelming amount of studies (e.g., Fazio & Olson, 2003; Herring et al., 2013; Klauer & Musch, 2003). Likewise, when a non-evaluative semantic judgment of the target stimuli is required, one can expect participants to process the primes in terms of that specific semantic stimulus dimension, thereby producing a non-evaluative semantic priming effect (e.g., Klinger, Burton, & Pitts, 2000).

In principle, however, the mechanism of task misapplication fails to account for priming effects that arise in the absence of overlap between the prime set and the response set. As an example, consider a pronunciation task in which participants are presented with positive and negative words as primes and targets and are asked to pronounce the targets. While task-misapplication in the pronunciation task may (perhaps) result in a tendency to pronounce the prime words, automatic evaluative stimulus processing is not expected to occur in this task because it is not required for the task at hand. Nevertheless, a number of researchers did report that they replicated the evaluative priming effect in the pronunciation task (Bargh, Chaiken, Raymond, & Hymes, 1996; Hermans, Houwer, & Eelen, 1994 Experiment 2; but see Klauer & Musch, 2001; Spruyt, Hermans, Pandelaere, De Houwer, & Eelen, 2004).

One way to reconcile these observations with the task-misapplication account is to assume that task-misapplication is not limited to proximal task goals (e.g., pronouncing words) but can also extend to more distal task goals (e.g., evaluating stimuli) that are relevant in a particular context (e.g., an evaluative priming experiment). In line with this reasoning, Spruyt and colleagues demonstrated that the evaluative priming effect replicates in the pronunciation task only if the experimental conditions encourage participants to engage in evaluative stimulus processing (Everaert, Spruyt, & De Houwer, 2011; Spruyt, 2014; Spruyt, De Houwer, Everaert, & Hermans, 2012; Spruyt, De Houwer, Hermans, & Eelen, 2007; Spruyt, Houwer, & Hermans, 2009; Spruyt & Tibboel, 2015). As an example, consider the findings of Spruyt et al. (2009). In their Experiment 3, participants were asked to pronounce target words on just 25 % of the trials. On the remaining trials, one group of participants was asked to categorize the targets in terms of their valence (i.e., the evaluative condition) whereas a second group was asked to indicate whether the target words referred to persons or objects (i.e., the non-evaluative condition). Crucially, the prime set and the target set used for the pronunciation trials consisted of positive and negative words that referred to persons and objects. It was thus possible to study the evaluative priming effect and the (non-evaluative) semantic priming effect at the same time. Results showed that the evaluative priming effect on the pronunciation trials was reliable in the evaluative condition but not in the non-evaluative condition. Conversely, the non-evaluative priming effect on the pronunciation trials was reliable in the non-evaluative condition but not in the evaluative condition. Clearly, these findings can be readily accounted for if one assumes that participants misapplied the (dominant) categorization task to the primes presented on the critical pronunciation trials.

Interestingly, this reasoning can also account for the sequential priming data reported by Pecher, Zeelenberg, and Raaijmakers (1998). They examined whether target responding in the pronunciation task and the lexical decision task can be affected by the degree of perceptual relatedness between the primes and the targets. To that end, participants were presented with primes and targets that referred to objects with a similar shape (e.g., pizza – coin) or a dissimilar shape (e.g., pen – coin). As expected, response latencies were found to be faster on congruent trials as compared to incongruent trials, but only if participants completed a shape-categorization task prior to the experimental priming task. It could thus be argued that the perceptual priming effects reported by Pecher et al. (1998) resulted from the fact that participants misapplied the shape-categorization task to the primes presented during the sequential priming procedure, thereby producing a perceptual priming effect.

In sum, the idea that task-misapplication can occur for proximal as well as distal task sets broadens the scope of our findings dramatically. Still, it must be reiterated that the emergence of a sequential priming effects requires two, sequential processes (Spruyt et al., 2011). While our work clearly shows that task-misapplication drives the interpretation of stimuli that are presented under automaticity conditions (see also Ansorge et al., 2014), we make no general claims about the mechanisms that may or may not be responsible for translating the outcome of the task-misapplication process into an observable sequential priming effect.

Conclusion

While it is typically argued that non-automatic processes are required to process the relation of one stimulus relative to other stimuli, we presented strong evidence for the hypothesis that complex relational stimulus processing can take place under automaticity conditions. In addition, our findings strongly suggest that automatic relational stimulus processing is driven by an unintentional misapplication of (proximal and/or distal) task sets, a viewpoint that can explain in a parsimonious way various, often conflicting results that have been reported in the literature.

REFERENCES

Ansorge, U., & Heumann, M. (2003). Top-down contingencies in peripheral cuing: The roles of color and location. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 937–948. doi: 10.1037/0096-1523.29.5.937

Ansorge, U., Heumann, M., & Scharlau, I. (2002). Influences of visibility, intentions, and probability in a peripheral cuing task. *Consciousness and Cognition*, *11*, 528–545. doi: 10.1016/S1053-8100(02)00026-0

Ansorge, U., Kiefer, M., Khalid, S., Grassl, S., & König, P. (2010). Testing the theory of embodied cognition with subliminal words. *Cognition*, *116*, 303–320. doi: 10.1016/j.cognition.2010.05.010

Ansorge, U., Kiss, M., & Eimer, M. (2009). Goal-driven attentional capture by invisible colors: Evidence from event-related potentials. *Psychonomic Bulletin & Review*, *16*, 648–653. doi: 10.3758/PBR.16.4.648

Ansorge, U., Kunde, W., & Kiefer, M. (2014). Unconscious vision and executive control: How unconscious processing and conscious action control interact. *Consciousness and Cognition*, *27*, 268–287. doi: 10.1016/j.concog.2014.05.009

Ansorge, U., & Neumann, O. (2005). Intentions Determine the Effect of Invisible Metacontrast-Masked Primes: Evidence for Top-Down Contingencies in a Peripheral Cuing Task. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 762–777. doi: 10.1037/0096-1523.31.4.762

Banks, W. P. (1977). Encoding and processing of symbolic information in comparative judgments. In G. H. Bower (Ed.), *The Psychology of learning and motivation: Advantages in research and theory* (Vol. 11, pp. 101–159). New York, NY: Academic Press.

Barcelo, F., Escera, C., Corral, M. J., & Periáñez, J. A. (2006). Task Switching and Novelty Processing Activate a Common Neural Network for Cognitive Control. *Journal of Cognitive Neuroscience*, *18*, 1734–1748. doi: 10.1162/jocn.2006.18.10.1734

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The Automatic Evaluation Effect: Unconditional Automatic Attitude Activation with a Pronunciation Task. *Journal of Experimental Social Psychology*, *32*, 104–128. doi: 10.1006/jesp.1996.0005

Bargh, J. A., Schwader, K. L., Hailey, S. E., Dyer, R. L., & Boothby, E. J. (2012). Automaticity in social-cognitive processes. *Trends in Cognitive Sciences*, *16*, 593–605. doi: 10.1016/j.tics.2012.10.002

Damian, M. F. (2001). Congruity effects evoked by subliminally presented primes: Automaticity rather than semantic processing. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 154–165. doi: 10.1037//0096-1523.27.1.154

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., … Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, *395*, 597–600. doi: 10.1038/26967

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368. doi: 10.1037/a0014211

Dijksterhuis, A., Smith, P. K., van Baaren, R. B., & Wigboldus, D. H. J. (2005). The Unconscious Consumer: Effects of Environment on Consumer Behavior. *Journal of Consumer Psychology*, *15*, 193–202. doi: 10.1207/s15327663jcp1503_3

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience*, *3*, 277–283. doi: 10.1038/72991

Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General*, *127*, 286–303. doi: 10.1037/0096-3445.127.3.286

Everaert, T., Spruyt, A., & De Houwer, J. (2011). On the (un)conditionality of automatic attitude activation: The valence proportion effect. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *65*, 125–132. doi: 10.1037/a0022316

Fazio, R. H., & Olson, M. A. (2003). Implicit Measures in Social Cognition Research: Their Meaning and Use. *Annual Review of Psychology*, *54*, 297–327. doi: 10.1146/annurev.psych.54.101601.145225

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238. doi: 10.1037/0022-3514.50.2.229

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.

Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2002). *Relational Frame Theory - A Post-Skinnerian Account of Human Language and Cognition*. Boston, MA: Kluwer Academic Publishers.

Hermans, D., Crombez, G., & Eelen, P. (2000). Automatic attitude activation and efficiency: The fourth horseman of automaticity. *Psychologica Belgica*.

Hermans, D., Houwer, J. D., & Eelen, P. (1994). The affective priming effect: Automatic activation of evaluative information in memory. *Cognition & Emotion*, *8*, 515–533. doi: 10.1080/02699939408408957

Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., … Crites, S. L. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychological Bulletin*, *139*, 1062–1089. doi: 10.1037/a0031309

Kiefer, M. (2002). The N400 is modulated by unconsciously perceived masked words: further evidence for an automatic spreading activation account of N400 priming effects. *Cognitive Brain Research*, *13*, 27–39. doi: 10.1016/S0926-6410(01)00085-4

Kiefer, M., & Martens, U. (2010). Attentional sensitization of unconscious cognition: Task sets modulate subsequent masked semantic priming. *Journal of Experimental Psychology: General*, *139*, 464–489. doi: 10.1037/a0019561

Kiefer, M., Sim, E.-J., & Wentura, D. (2015). Boundary conditions for the influence of unfamiliar non-target primes in unconscious evaluative priming: The moderating role of attentional task sets. *Consciousness and Cognition*, 342–356. doi: 10.1016/j.concog.2015.01.010

Klauer, K. C., Draine, S. C., & Greenwald, A. G. (1998). An unbiased errors-in-variables approach to detecting unconscious cognition. *British Journal of Mathematical and Statistical Psychology*, *51*, 253–267. doi: 10.1111/j.2044-8317.1998.tb00680.x

Klauer, K. C., Greenwald, A. G., & Draine, S. C. (1998). Correcting for measurement error in detecting unconscious cognition: Comment on Draine and Greenwald (1998). *Journal of Experimental Psychology: General*, *127*, 318–319. doi: 10.1037/0096-3445.127.3.318

Klauer, K. C., & Musch, J. (2001). Does sunshine prime loyal? Affective priming in the naming task. *The Quarterly Journal of Experimental Psychology Section A*, *54*, 727–751. doi: 10.1080/713755986

Klauer, K. C., & Musch, J. (2003). Affective priming: Findings and theories. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 7–49). Mahwah, NJ: Erlbaum.

Klauer, K. C., & Teige-Mocigemba, S. (2007). Controllability and resource dependence in automatic evaluation. *Journal of Experimental Social Psychology*, *43*, 648–655. doi: 10.1016/j.jesp.2006.06.003

Klinger, M. R., Burton, P. C., & Pitts, G. S. (2000). Mechanisms of unconscious priming: I. Response competition, not spreading activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 441–455. doi: 10.1037/0278-7393.26.2.441

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility--A model and taxonomy. *Psychological Review*, *97*, 253–270. doi: 10.1037/0033-295X.97.2.253

Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition*, *88*, 223–242. doi: 10.1016/S0010-0277(03)00023-4

Kutner, M., Nachtsheim, C., Neter, J., & Li, W. (2005). *Applied linear statistical models*. Europe: McGraw-Hill Irwing.

Moors, A., & De Houwer, J. (2001). Automatic appraisal of motivational valence: Motivational affective priming and Simon effects. *Cognition & Emotion*, *15*, 749–766. doi: 10.1080/02699930143000293

Moors, A., & De Houwer, J. (2006). Automaticity: A Theoretical and Conceptual Analysis. *Psychological Bulletin*, *132*, 297–326. doi: 10.1037/0033-2909.132.2.297

Moors, A., De Houwer, J., & Eelen, P. (2004). Automatic stimulus-goal comparisons: Support from motivational affective priming studies. *Cognition & Emotion*, *18*, 29–54. doi: 10.1080/02699930244000462

Moors, A., De Houwer, J., Hermans, D., & Eelen, P. (2005). Unintentional processing of motivational valence. *The Quarterly Journal of Experimental Psychology Section A*, *58*, 1043–1063. doi: 10.1080/02724980443000467

Moors, A., Spruyt, A., & De Houwer, J. (2010). In Search of a Measure That Qualifies as Implicit. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 19–37). New York, NY: Guilford Press.

Naccache, L., & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition*, *80*, 215–229. doi: 10.1016/S0010-0277(00)00139-6

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: from measures to mechanisms. *Trends in Cognitive Sciences*, *15*, 152–159. doi: 10.1016/j.tics.2011.01.005

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*, 107–123. doi: 10.1016/S0042-6989(01)00250-4

Pecher, D., Zeelenberg, R., & Raaijmakers, J. G. W. (1998). Does Pizza Prime Coin? Perceptual Priming in Lexical Decision and Pronunciation. *Journal of Memory and Language*, *38*, 401–418. doi: 10.1006/jmla.1997.2557

Quinn, W. M., & Kinoshita, S. (2008). Congruence effect in semantic categorization with masked primes with narrow and broad categories. *Journal of Memory and Language*, *58*, 286–306. doi: 10.1016/j.jml.2007.03.004

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*, 510–532. doi: 10.1037/0033-2909.114.3.510

Reynvoet, B., Gevers, W., & Caessens, B. (2005). Unconscious Primes Activate Motor Codes Through Semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 991–1000. doi: 10.1037/0278-7393.31.5.991

Schneider, D. W., & Logan, G. D. (2007). Defining task-set reconfiguration: The case of reference point switching. *Psychonomic Bulletin & Review*, *14*, 118–125. doi: 10.3758/BF03194038

Severens, E., Lommel, S. V., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, *119*, 159–187. doi: 10.1016/j.actpsy.2005.01.002

Simonson, I. (1989). Choice Based on Reasons: The Case of Attraction and Compromise Effects. *Journal of Consumer Research*, *16*, 158–174. doi: 10.1086/209205

Spruyt, A. (2014). Attention please: Evaluative priming effects in a valent/non-valent categorisation task (reply to Werner & Rothermund, 2013). *Cognition & Emotion*, *28*, 560–569. doi: 10.1080/02699931.2013.832653

Spruyt, A., Clarysse, J., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2010). Affect 4.0: A Free Software Package for Implementing Psychological and Psychophysiological Experiments. *Experimental Psychology (formerly Zeitschrift Für Experimentelle Psychologie)*, *57*, 36–45. doi: 10.1027/1618-3169/a000005

Spruyt, A., De Houwer, J., Everaert, T., & Hermans, D. (2012). Unconscious semantic activation depends on feature-specific attention allocation. *Cognition*, *122*, 91–95. doi: 10.1016/j.cognition.2011.08.017

Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective Priming of Nonaffective Semantic Categorization Responses. *Experimental Psychology (formerly "Zeitschrift Für Experimentelle Psychologie")*, *54*, 44–53. doi: 10.1027/1618-3169.54.1.44

Spruyt, A., Gast, A., & Moors, A. (2011). The sequential priming paradigm: A primer. In K. C. Klauer, C. Stahl, & A. Voss (Eds.), *Cognitive methods in social psychology* (pp. 48–77). New York, NY: Guilford Press.

Spruyt, A., Hermans, D., Houwer, J. D., & Eelen, P. (2002). On The Nature of the Affective Priming Effect: Affective Priming of Naming Responses. *Social Cognition*, *20*, 227–256. doi: 10.1521/soco.20.3.227.21106

Spruyt, A., Hermans, D., Pandelaere, M., De Houwer, J., & Eelen, P. (2004). On the Replicability of the Affective Priming Effect in the Pronunciation Task. *Experimental Psychology (formerly "Zeitschrift Für Experimentelle Psychologie")*, *51*, 109–115. doi: 10.1027/1618-3169.51.2.109

Spruyt, A., Houwer, J. D., & Hermans, D. (2009). Modulation of automatic semantic priming by feature-specific attention allocation. *Journal of Memory and Language*, *61*, 37–54. doi: 10.1016/j.jml.2009.03.004

Spruyt, A., & Tibboel, H. (2015). On the Automaticity of the Evaluative Priming Effect in the Valent/Non-Valent Categorization Task. *PLOS ONE*, *10*, e0121564. doi: 10.1371/journal.pone.0121564

Ulrich, M., Hoenig, K., Grön, G., & Kiefer, M. (2013). Brain Activation during Masked and Unmasked Semantic Priming: Commonalities and Differences. *Journal of Cognitive Neuroscience*, *25*, 2216–2229. doi: 10.1162/jocn_a_00449

Van Opstal, F., Gevers, W., Osman, M., & Verguts, T. (2010). Unconscious task application. *Consciousness and Cognition*, *19*, 999–1006. doi: 10.1016/j.concog.2010.05.002

Wentura, D., & Frings, C. (2008). Response-bound primes diminish affective priming in the naming task. *Cognition & Emotion*, *22*, 374–384. doi: 10.1080/02699930701446064

Wiers, R. W., Van Woerden, N., Smulders, F. T. Y., & De Jong, P. J. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology*, *111*, 648–658. doi: 10.1037/0021-843X.111.4.648

CHAPTER 3

ON THE RELATIONSHIP BETWEEN BODY DISSATISFACTION AND IMPLICIT BELIEFS ABOUT ACTUAL AND IDEAL BODY IMAGE [1]

ABSTRACT

We examined whether implicit measures of actual and ideal body image can be used to predict body dissatisfaction in young female adults. Participants completed two Implicit Relational Assessment Procedures (IRAPs) to examine implicit beliefs concerning their actual (e.g., I *am* thin) and ideal body image (e.g., I *want to be* thin). Body dissatisfaction was examined via self-report questionnaires and rating scales. As expected, the implicit belief to be thin (i.e., actual body image) was lower in participants who exhibited a high degree of body dissatisfaction than in participants who exhibited a low degree of body dissatisfaction. In contrast, the implicit desire to be thin (i.e., ideal body image) was stronger in participants who exhibited a high level of body dissatisfaction than in participants who were less dissatisfied with their body. Adding further weight to the idea that both IRAP measures captured different underlying constructs, we also observed that they correlated differently with body mass index, explicit body dissatisfaction, and explicit measures of actual and ideal body image. More generally, these findings underscore the advantage of using implicit measures that incorporate relational information relative to implicit measures that allow for an assessment of associative relations only.

[1] Based on Heider, N., Spruyt, A., & De Houwer, J. (2015). On the Relationship between Body Dissatisfaction and Implicit Beliefs about Actual and Ideal Body Image. *Manuscript submitted for publication.*

INTRODUCTION

Body dissatisfaction can be defined as the negative attitude towards one's own body resulting from a (perceived) discrepancy between one's actual physical appearance (i.e., the actual body image) and internalized ideals about one's physical appearance (i.e., the ideal body image; e.g., Cash & Szymanski, 1995; Higgins, 1989; Strauman, Vookles, Berenstein, Chaiken, & Higgins, 1991; Williamson, Gleaves, Watkins, & Schlundt, 1993). Because body dissatisfaction plays a central role in the causation and maintenance of eating disorders (American Psychiatric Association, 2013; Fairburn & Harrison, 2003; Stice, 2001), behavioral scientists have long sought ways to measure the degree of dissatisfaction with one's physical appearance. Most often, they rely on the use of direct self-report measures (i.e., questionnaires), but it is well-known that such measures can be susceptible to social desirability and impression management (Cronbach, 1990; Holtgraves, 2004). In addition, self-report measures are, by definition, unsuited to capture attitudes that are introspectively unidentified (Greenwald & Banaji, 1995). Accordingly, behavioral scientists have begun developing diagnostic instruments that allow for an assessment of body dissatisfaction in an indirect way, that is, without having to ask for a direct self-assessment. Instead, inter-individual differences are inferred from a respondent's response pattern in well-controlled computer tasks, often referred to as *implicit measures* (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009).

Hitherto, several attempts have been undertaken to develop implicit measures of body-related attitudes (e.g., Ahern, Bennett, & Hetherington, 2008; Bessenoff & Sherman, 2000; Bluemke & Friese, 2012; Degner & Wentura, 2009; Juarascio et al., 2011; Parling, Cernvall, Stewart, Barnes-Holmes, & Ghaderi, 2012; Roddy, Stewart, & Barnes-Holmes, 2010, 2011; Watts, Cranney, & Gleitzman, 2008). Consider, for example, the findings of Bluemke and Friese (2012). In their adaptation of the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), stimuli referring to *thinness* and *overweight* and stimuli related to either *the self* or related to *a well-known other person* (i.e., a friend or a relative) were presented one by one on a computer screen. In a first critical block of trials, participants were asked to press one key as quickly as possible upon the presentation of a word referring to thinness (e.g., 'skinny') or a word related to the self (e.g., the participant's first name). The second key was to be pressed upon the presentation of a word referring to overweight (e.g., 'fat') or a word related to another person (e.g., the first name of a friend). In a second critical block, response assignments were reversed so that words referring to thinness and another person were assigned to the first key whereas words referring to overweight and the self were assigned to the second key. Based on the assumption that it is easier to respond when concepts assigned to the same key are associated in memory, a person's implicit body image was then inferred from the difference in performance between the two critical blocks. More specifically, participants with a thin body image were expected to perform best when words relating to the self and thinness were assigned to the same key whereas participants with an overweight body image were expected to perform best when words relating to the self and overweight were assigned to the same key. In sum, the IAT measure developed by Bluemke and Friese (2012) was designed to capture inter-individual differences in the strength of associations between the concepts *self* and *body size* (i.e., thinness/overweight).

Body dissatisfaction, however, by definition comprises more than a simple association between the self and body size because it is driven by the (perceived) discrepancy between one's ideal and actual body image. Crucially, both ideal and actual body image involve a relation between the concepts *self* and *body size* but differ with regard to how those concepts are related. More specifically, beliefs about actual body image are characterized by a descriptive relation (e.g., I *am* thin). In contrast, beliefs about ideal body image relate the self to body size in terms of desirability (e.g., I *want to be* thin). Clearly, these two beliefs are fundamentally different, yet the IAT measure of Bluemke and Friese (2012) is unable to differentiate between them as the IAT was designed to capture the associative, unqualified strength between two concepts.

As pointed out by Hughes, Barnes-Holmes, and Vahey (2012; also see De Houwer, 2014), this is a property not only of the IAT but also of several other implicit measures such as the (standard) evaluative priming task (EPT; Fazio, Jackson, Dunton, & Williams, 1995) and the affect misattribution paradigm (AMP; Payne, Cheng, Govorun, & Stewart, 2005).[2] However, the Implicit Relational Assessment Procedure (IRAP), developed by Barnes-Holmes et al. (2006), was designed specifically to capture inter-individual differences in the extent to which respondents relate concepts in a specific manner (Hughes & Barnes-Holmes, 2013). It could thus be hypothesized that the IRAP has the potential to outperform classic implicit measures such as the EPT and the IAT whenever relational information is critical.

As an example, consider the studies of Remue and colleagues who examined ideal and actual self-esteem in a sample of dysphoric students and a non-dysphoric control group (Remue, De Houwer, Barnes-Holmes, Vanderhasselt, & De Raedt, 2013; Remue, Hughes, De Houwer, & De Raedt, 2014). To capture ideal self-esteem, participants were presented, on each of a series of trials, with either the stimulus 'I want to be' or 'I do not want to be' together with either a positive or negative adjective. Each combination of both stimuli was thus either congruent or incongruent with having positive ideal self-esteem (e.g., 'I want to be + good' and 'I want to be + bad', respectively). Crucially, in one block of trials, participants were asked to respond as if they had the desire to possess positive self-esteem. They were thus required to select the response 'true' whenever a stimulus combination was presented that referred to having positive ideal self-esteem (e.g., 'I do not want to be + bad'). Conversely, whenever a stimulus combination was presented that referred to having negative ideal self-esteem (e.g., 'I want to be + bad'), they were required to select the response 'false'. In a second block of trials, participants were asked to respond as if they had the desire to possess negative self-esteem. Accordingly, they were expected to select the response 'true' whenever a stimulus combination was presented that referred to having negative ideal self-esteem (e.g., 'I want to be + bad'). Whenever a stimulus combination was presented that referred to having positive ideal self-esteem (e.g., 'I want to be + good'), they were expected to select the response 'false'. The IRAP is based on the assumption that it is more difficult to respond in a manner that is inconsistent with one's personal beliefs than it is to respond in a manner that is consistent with one's personal beliefs. A person's level of ideal self-esteem was thus inferred from the difference in performance

[2] It might be noted that the sequential priming paradigm can be adapted so as to allow for the measurement of (automatic) relational processing. For an example, see Heider, Spruyt, and De Houwer (2015).

between the two critical blocks. More specifically, participants with positive ideal self-esteem were expected to perform best when responding to statements that were in line with desiring to have positive self-esteem whereas participants with negative ideal self-esteem were expected to perform best when responding to statements in line with desiring to have negative self-esteem. It was anticipated that the desire to be good would be more pronounced in dysphoric students than in non-dysphoric controls, as was indeed observed by Remue et al. (2013, but see Remue et al., 2014).

Crucially, Remue and colleagues also administered a second IRAP that was designed to capture actual self-esteem. This second IRAP was identical to the first one, except for the fact that participants were now (a) presented with stimulus combinations that referred to actual self-esteem (e.g., 'I am + good') and (b) were required to respond as if they did or did not possess positive self-esteem. Despite the structural similarity between the two IRAPs, the results obtained with the second IRAP revealed no (Remue et al., 2014) or even a reversed difference (i.e., more positive actual self-esteem in non-dysphoric than in dysphoric students; Remue et al., 2013) between dysphoric and non-dysphoric students. These findings are in line with the notion that ideal and actual self-esteem are two different constructs and should therefore be measured independently from each other. More generally, because the essential difference between ideal and actual self-esteem concerns the quality of the relation between the self and positive/negative affect, these findings underscore the need for implicit measures that are sensitive to relational information.

Accordingly, the aim of the present research was to develop an implicit measure of body dissatisfaction that takes into account the way in which the concepts self and body-size are related. Participants were asked to perform two IRAPs in quick succession, one to capture the extent to which participants endorsed or rejected beliefs reflecting their actual body (e.g., 'I am thin', 'I am not overweight') and one to capture the extent to which participants endorsed or rejected beliefs reflecting their ideal body (e.g., 'I want to be thin', 'I don't want to be overweight'). We expected that the discrepancy between the two IRAP scores would depend upon the degree of self-reported body dissatisfaction. Specifically, we expected that the belief to be thin would be more pronounced in participants low in body dissatisfaction as compared to participants high in body dissatisfaction. In contrast, we expected that the desire to be thin would be less pronounced in participants low in body dissatisfaction as compared to participants high in body dissatisfaction.

METHOD

Participants

About one month prior to the actual experiment, 307 students at Ghent University completed the body dissatisfaction subscale of the Eating Disorders Inventory (EDI; Garner, Olmstead, & Polivy, 1983) during an online screening study that involved several questionnaires. Next, using an online participant recruitment system, we invited all female students who had scored within the first and forth quartile of the total EDI distribution (N = 112). In total, 52 female students (M = 19.6 years, SD = 4.8) participated in the actual experiment in exchange for course credit. Four participants were excluded because they failed to complete at least one of the two IRAP measures. Based on the scores of two body dissatisfaction self-reports obtained during the actual test session, the final sample (N = 48) was divided into a low body dissatisfaction group (n = 24, M = -0.90, SD = 0.27, min = -1.17, max = -0.03) and a high body dissatisfaction group (n = 24, M = 0.90, SD = 0.43, min = 0.04, max = 1.75), $t(46) = 17.13$, $p < .001$, d = 4.95).[3] All participants were Dutch speakers and had normal or corrected-to-normal vision.

Measures

Self-report measures. Body dissatisfaction was assessed by means of the body dissatisfaction subscale of the Eating Disorder Inventory (EDI, 9 items; Garner et al., 1983) as well as the body dissatisfaction subscale of the Body Attitude Test (BAT, 7 items; Probst, Vandereycken, Coppenolle, & Vanderlinden, 1995). Scores of the two measures were highly correlated, r = .91. Scores of explicit body dissatisfaction were

[3] A renewed administration of explicit measures of body dissatisfaction during the test session was necessary as invitations were sent out anonymously via an online participant recruiting system.

computed for each participant by averaging their standardized scores of both body dissatisfaction self-report measures. Based on this general score of body dissatisfaction, participants were assigned into either the low or high body dissatisfaction group. Actual and ideal body image were measured using the female version of the Contour Drawing Rating Scale (CDRS; Thompson & Gray, 1995). The CDRS consists of nine schematic (female) figures of varying sizes ranging from *underweight* (1) to *overweight* (9). Participants completed the CDRS twice, once to indicate their actual body image and once to indicate their ideal body image. Higher CDRS scores indicated a more overweight actual and ideal body image, respectively. Finally, we computed two Body Mass Indices for each participant, once using their self-reported weight and height (i.e., self-reported BMI) and once using their factual weight (i.e., factual BMI). Discrepancies between self-reported BMI and factual BMI were unrelated to all other (indirect and direct) measures. Accordingly, only the factual BMI data were used.

IRAPs. Participants completed two IRAPs, one to capture actual body image (i.e., actual-IRAP) and one to capture ideal body image (i.e., ideal-IRAP). To capture actual body image, participants were presented with combinations of the stimuli ‘I am’ or ‘I am not’ (in the IRAP literature referred to as *sample stimuli*) and one of twelve words referring to the concepts thinness and overweight (in the IRAP literature referred to as *target stimuli*). All target stimuli are presented in the Appendix. The combination of sample and target stimuli resulted in two trial types: Twelve combinations were in line with the belief ‘I am thin’ (e.g., ‘I am + slim’; ’I am not + chubby’), and 12 combinations were in line with the belief ‘I am overweight’ (e.g., ‘I am + chubby’; ‘I am not + slim’). Similarly, to capture ideal body image, the same set of 12 target stimuli was combined with the sample stimuli ‘I want to be’ or ‘I don’t want to be’. Trials were thus either in line with the belief ‘I want to be thin’ (e.g., ‘I want to be + slim’; ’I don’t want to be + chubby’) or in line with the belief ‘I want to be overweight’ (‘I want to be + chubby’; ‘I don’t want to be + slim’).

Both IRAPs consisted of six blocks of trials in which each of 24 combinations of sample and target stimuli was presented exactly once in random order (144 trials in total). Participants were asked to respond as fast as possible by pressing one of two response keys. One of the keys indicated ‘true’ whereas the other indicated ‘false’.

Response assignments varied randomly from trial to trial. To capture actual body image, participants were asked to respond in line with the belief ‘I am thin’ in one type of block (i.e., congruent block). In the second type of block (i.e., incongruent block), they were asked to respond in line with the belief ‘I am overweight’. Likewise, to capture ideal body image, participants were asked to respond in line with the belief ‘I want to be thin’ in the congruent block. In the incongruent block, they were asked to respond in line with the belief ‘I want to be overweight’. Congruent and incongruent blocks were presented in an alternating order and all participants started with a congruent block of trials.

Each IRAP was preceded by a practice phase. All participants completed a congruent practice block followed by an incongruent practice block in which each of the 24 stimulus combinations were again presented in a random order. Participants were instructed verbally to focus on response accuracy first and then to increase their response speed. Unless participants achieved an accuracy of more than 80% and a median response latency of less than 2000 ms in both practice blocks, a second pair of practice blocks was presented. If necessary, this procedure was repeated after the second pair of practice trials. If participants failed to reach the threshold criteria during the third pair of practice trials, the IRAP was stopped. This was the case for four participants who did not complete at least one of the IRAPs.[4]

Each trial was started with a 400-ms blank white screen. Afterwards, one of the 24 stimulus combinations of a sample and a target stimulus was presented in two rows at the center of the upper half of the computer screen in black color, Arial font size 26. To signal the response assignments for each trial (i.e., the keys D and K), the words ‘true’ and ‘false’ were presented in a green color, Arial font size 36, at the bottom left and the bottom right corner of the computer screen. In case of an incorrect response, a red X was presented in Arial font size 48 below the stimulus combination. Participants were required to correct an erroneous response in order to proceed to the next trial. After each block, participants were presented with feedback about their accuracy (in

[4] Note, that the version of the IRAP software used in this study recorded the raw response latencies of the last pair of practice blocks only. Consequentially, there is no information about the number of participants that completed one, two, or three pairs of practice blocks, respectively, before proceeding with the test phase.

percentages) and speed of responding (in median response latencies), based on the last 24 trials, in red color, Arial font size 12. Each block of trials was preceded by the presentation of a reminder about the overall response rule, Arial font size 14 (e.g., 'Please respond as if you are thin and are not overweight'). All stimuli were presented on a 19-inch VGA screen (75Hz, 1024 × 768 pixels) and implemented using the 2012 version of the IRAP software downloaded from http://www.irapresearch.org.

Procedure

All participants gave informed consent before participation. Participants completed both IRAPs in a counterbalanced order. They then completed the EDI, BAT, and CDRSs for actual and ideal body image, in that order. Finally, the participants' weight and height were assessed for the calculation of the BMI. All participants were tested individually. The experiment took approximately 45 minutes.

RESULTS

Data Preparation

For each participant and each version of the IRAP, the raw response latencies of the six experimental blocks were aggregated in a single overall D-score using the algorithm described by Greenwald, Nosek, and Banaji (2003). We first excluded response latencies above 10,000 ms (0.10 %). Other criteria for data exclusion as specified in the scoring algorithm (e.g., responses faster than 300 ms on more than 10% of the trials) were not met. We then calculated standard deviations on the basis of all response latencies observed in subsequent pairs of congruent and incongruent blocks (i.e., Block 1 & 2, 3 & 4, and 5 & 6). Next, mean response latencies were calculated for each of the six blocks. Difference scores were then calculated for each pair of blocks by subtracting the mean response latency observed in the congruent block from the mean response latency observed in the incongruent block. Finally each difference score was divided by its corresponding standard deviation, yielding three D-scores, one for each pair of blocks.

Finally, the three D-scores were averaged to obtain one overall D-score.[5] D-scores were calculated so that positive values indicate a higher degree of a thin body image belief (actual or ideal).

Effects at the Group Level

To investigate whether implicit measures of actual and ideal body image were dependent upon the degree of body dissatisfaction, the overall D-scores were submitted to a 2 (*IRAP*: actual vs. ideal) x 2 (*body dissatisfaction*: high vs. low) mixed models ANOVA.[6] As expected, we found a significant interaction between body dissatisfaction and IRAP, $F(1, 46) = 6.71$, $p = .013$, $\eta_\rho^2 = .13$, indicating that the two groups responded differently to the two IRAPs. Participants low in body dissatisfaction scored higher on the actual-IRAP than participants high in body dissatisfaction, 0.13 ($SD = 0.17$) vs. 0.05 ($SD = 0.16$), respectively. In contrast, participants high in body dissatisfaction scored higher on the ideal-IRAP than participants low in body dissatisfaction, 0.13 ($SD = 0.16$) vs. 0.03 ($SD = 0.17$), respectively. This pattern of results is consistent with our hypotheses (a) that the implicit belief to be thin is more pronounced in individuals who are low in body dissatisfaction as compared to individuals who exhibit a high degree of body dissatisfaction, and (b) that the implicit desire to be thin is more pronounced in individuals who are high in body dissatisfaction as compared to individuals who exhibit a low degree of body dissatisfaction. Additional *t*-tests showed, however, that the D-scores of the actual-IRAP did not differ significantly between groups, $t(46) = 1.57$, $p = .122$, d = 0.45. In contrast, the D-scores of the ideal-IRAP did differ significantly between the groups, $t(46) = 2.12$, $p = .040$, d = 0.61.

[5] It may be noted that this approach diverges from the analysis strategy adopted by Barnes-Holmes, Barnes-Holmes, Stewart, and Boles (2011). Instead of calculating a single, overall D-score, these authors suggested to calculate four different D-scores, one for each of type of trial (i.e., 'I am + thin', 'I am not + thin', 'I am + overweight', 'I am not + overweight', in case of the actual-IRAP). The rationale to depart from this approach is threefold. First, for both of the IRAPs that we used, all four trial types were assumed the probe the exact same belief. Second, by aggregating all observations into a single D-score, the impact of outlying observations is much better controlled for. Finally, a single D-score can be interpreted in an intuitive manner.

[6] As explained above, the order in which the two IRAP tasks were completed was counterbalanced across participants. Reassuringly, our critical effect (i.e., the interaction between explicit body dissatisfaction and the type of IRAP) was not moderated by this counterbalancing factor, $F(1, 44) = 1.38$, $p = .247$, $\eta\rho^2 = .03$, nor were any other effects, all *F*s < 1, all *p*s > .49. We therefore excluded the order factor from the analyses.

Differences between actual and ideal body image were significant for participants low in body dissatisfaction, $t(46) = 2.02$, $p = .049$, d = 0.58, and marginally significant for participants high in body dissatisfaction, $t(46) = 1.65$, $p = .105$, d = 0.48. The ANOVA did not reveal other significant effects, all Fs < 1, all ps > .712.

Correlational Analyses

For exploratory reasons, we also computed all pairwise correlations between the IRAP scores, the CDRS measures, the explicit measure of body dissatisfaction, and the BMI (see Table 1). Note, however, that most variables were not normally distributed because our sample consisted of participants who exhibited either a high or low degree of body dissatisfaction. Accordingly, Spearman's rank order correlation coefficients were computed rather than Pearson product-moment correlation coefficients. In line with the results presented above, the ideal-IRAP correlated or tended to correlate positively with the actual-CDRS, the ideal-CDRS, the explicit measure of body dissatisfaction, and the BMI. Conversely, the actual-IRAP correlated negatively with each of these measures, albeit not significantly so. Adding further weight to the idea that both IRAP measures captured different underlying constructs, the degree to which the two IRAP measures correlated with each of the other variables was reliably different for the actual-CDRS, the ideal-CDRS, and the explicit measure of body dissatisfaction, ts < 2.01, ps < .05. For the BMI, the difference between the correlation with the ideal-IRAP and the actual-IRAP just missed significance, $t(46) = 1.91$, $p = .06$.

Table 1. *Correlations between measures.*

		1	2	3	4	5	6
1	Actual-IRAP	-	-0.08	-0.24^{+}	-0.27^{+}	-0.12	-0.12
2	Ideal-IRAP		-	0.39^{**}	0.27^{+}	0.30^{*}	0.28^{*}
3	Actual-CDRS			-	0.60^{***}	0.82^{***}	0.77^{***}
4	Ideal-CDRS				-	0.17	0.63^{***}
5	EBD					-	0.61^{***}
6	BMI						-

Note. IRAP = Implicit Relational Assessment Procedure; CDRS = Contour Drawing Rating Scale; EBD = explicit body dissatisfaction, based on averaged scores of the body dissatisfaction subscales of the Eating Disorders Inventory (EDI) and the Body Attitude Test (BAT); BMI = Body Mass Index

$^{+}$ p < .10, * p < .05, ** p < .01, *** p < .001

Reliability

For the EDI and BAT subscales of body dissatisfaction, Chronbach's alpha was 0.99 and 0.93, respectively. Test-retest reliability of the actual body image CDRS was r = .78, as reported by Thompson & Gray (1995). Reliability coefficients of both IRAPs were estimated using a bootstrap procedure, wherein 100 random-splits were drawn from the data. The calculation steps for the random-split D-scores were identical to the calculation of the overall D-scores as presented above. For each random split, the correlation across participants between the two D-scores was calculated. Correlations were then averaged. This procedure resulted in spearman-brown corrected mean split-half correlations of Rsb = 0.32 and Rsb = 0.42, for the actual and ideal body image IRAP, respectively.

DISCUSSION

The degree to which people are dissatisfied with their own body is a function of the (perceived) discrepancy between one's actual and ideal body image (Cash & Szymanski, 1995; Higgins, 1989). We hypothesized that participants high and low in body dissatisfaction would differ not only in their self-reported degree of body dissatisfaction but also in their implicit beliefs concerning their actual and ideal body image. More specifically, we expected the implicit belief to be thin to be more pronounced for participants low in body dissatisfaction as compared to participants high in body dissatisfaction. In contrast, we expected the implicit desire to be thin to be more pronounced in participants who exhibit a high degree of body dissatisfaction as compared to participants low in body dissatisfaction. Using the IRAP (Barnes-Holmes et al., 2006) as an implicit measure of beliefs, we found strong supporting evidence for both predictions. In addition, the pattern of correlations between each of the two IRAP measures and a number of other variables was quite different. In line with the idea that the desire to be thin must be higher in individuals who estimate their own physical appearance to be overweight, the ideal-IRAP correlated positively with the actual-CDRS. Likewise, significant positive correlations were observed with the BMI and explicitly measured body dissatisfaction. In contrast all correlations between the actual-IRAP scores and each of these measures were negative, albeit not significantly so. Taken together, these findings strongly suggest that both IRAP measures, despite their structural similarity, captured different underlying constructs.

These observations are important for two reasons. First, as pointed out above, an important rationale for using implicit measures is their alleged resistance to effects of social desirability and impression management. In addition, it has been argued that implicit measures may be used to capture traces of prior experience that are introspectively unidentified. It can thus be hypothesized that the added value of using implicit measures of actual and ideal body image to predict behavioral outcomes will be most pronounced when participants are somehow unwilling or unable to complete explicit measures of body dissatisfaction in a truthful manner. In this respect, it seems particularly interesting to use the IRAP measures developed here in the context of eating disorders (e.g., anorexia nervosa), for two reasons. First, patients suffering from eating disorder might be inclined to fake self-report measures of (ideal and actual) body image because of significant therapeutic consequences (e.g., compulsory admission). In addition, the discrepancy between actual body weight (BMI) and implicit beliefs about one's actual body weight might be an important cognitive marker for future therapeutic outcomes. In sum, while the current study revealed no added value of the IRAP measures over and above explicit measures (e.g., the CDRS measure of actual body image) in predicting self-reported body dissatisfaction, there are good reasons to suspect that implicit measures of ideal and actual body image may be much more instrumental in applied research contexts.

A second reason why our findings are important concerns the use of implicit measures in general. For two decades now, implicit measures have been widely used in various research domains, including health and clinical psychology (e.g., Descheemaeker, Spruyt, & Hermans, 2014; Spruyt et al., 2013; Stacy & Wiers, 2010; Teachman, Cody, & Clerkin, 2010; Wiers, Houben, Roefs, De Jong, Peter, & Etchison, 2010; Wiers, Van Woerden, Smulders, & De Jong, 2002), forensic psychology (e.g., Snowden & Gray, 2010), and consumer psychology (e.g., Perkins & Forehand, 2010). Crucially, traditional implicit measures such as the IAT, EPT, and AMP were designed to capture the extent to which certain concepts are associated in memory (Hughes et al., 2012). Each of these measures, for example, can be readily used to capture the extent to which a certain class of stimuli (e.g., spiders) is associated with a positive or negative valence. In many cases, however, it is not only important to examine whether two concepts are related in memory but also the precise way in which they are related. As demonstrated by Remue et al. (2013, 2014), for example, it makes a tremendous difference to know whether someone has the implicit belief to be a person who actually *is* good or *wants to be* good. Likewise, the findings reported here demonstrate that it is important to distinguish between the implicit *desire to be* thin versus the implicit *belief to be* thin. Accordingly, it seems most interesting or even essential to invest in the development of implicit measures that are capable of tapping into relational information that is more complex than simple, unqualified associations (e.g., The Relational Responding Task, recently introduced by De Houwer, Heider, Spruyt, Roets, & Hughes, 2015, also see Chapter 4).

To sum up, we used two IRAP measures: One to capture implicit beliefs concerning one's actual body image and one to capture implicit beliefs concerning one's ideal body image. Both IRAP measures were related to different outcome variables to a different extent, thereby underscoring the validity of both measures. More generally, our findings indicate that it is key to examine not only whether two concepts are related in memory but also *how* they are related. Future research concerning the applied value of implicit measures would thus benefit greatly from taking into account complex relational information.

REFERENCES

Ahern, A. L., Bennett, K. M., & Hetherington, M. M. (2008). Internalization of the Ultra-Thin Ideal: Positive Implicit Associations with Underweight Fashion Models are Associated with Drive for Thinness in Young Women. *Eating Disorders*, *16*, 294–307. doi: 10.1080/10640260802115852

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition). Arlington, VA: American Psychiatric Pub.

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, *32*, 169–177.

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2011). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *The Psychological Record*, *60*, 527–542.

Bessenoff, G. R., & Sherman, J. W. (2000). Automatic and Controlled Components of Prejudice Toward Fat People: Evaluation Versus Stereotype Activation. *Social Cognition*, *18*, 329–353. doi: 10.1521/soco.2000.18.4.329

Bluemke, M., & Friese, M. (2012). On the Validity of Idiographic and Generic Self-Concept Implicit Association Tests: A Core-Concept Model. *European Journal of Personality*, *26*, 515–528. doi: 10.1002/per.850

Cash, T. F., & Szymanski, M. L. (1995). The Development and Validation of the Body-Image Ideals Questionnaire. *Journal of Personality Assessment*, *64*, 466–477. doi: 10.1207/s15327752jpa6403_6

Cronbach, L. (1990). *Essentials of psychological testing*. New York, NY: Harper & Row.

Degner, J., & Wentura, D. (2009). Not Everybody Likes the Thin and Despises the Fat: One's Weight Matters in the Automatic Activation of Weight-Related Social Evaluations. *Social Cognition*, *27*, 202–221. doi: 10.1521/soco.2009.27.2.202

De Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass*, *8*, 342–353. doi: 10.1111/spc3.12111

De Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The Relational Responding Task: Toward a New Implicit Measure of Beliefs. *Frontiers in Psychology*, *6:319*. doi: 10.3389/fpsyg.2015.00319

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368. doi: 10.1037/a0014211

Descheemaeker, M., Spruyt, A., & Hermans, D. (2014). On the relationship between the indirectly measured attitude towards beer and beer consumption: the role of attitude accessibility. *PloS One*, *9*, e95302. doi: 10.1371/journal.pone.0095302.t004

Fairburn, C. G., & Harrison, P. J. (2003). Eating disorders. *The Lancet*, *361*, 407–416. doi: 10.1016/S0140-6736(03)12378-1

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027. doi: 10.1037/0022-3514.69.6.1013

Garner, D. M., Olmstead, M. P., & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders*, *2*, 15–34. doi: 10.1002/1098-108X(198321)2:2<15::AID-EAT2260020203>3.0.CO;2-6

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27. doi: 10.1037/0033-295X.102.1.4

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216. doi: 10.1037/0022-3514.85.2.197

Heider, N., Spruyt, A., & De Houwer, J. (2015). On the Automaticity of Relational Stimulus Processing. *Manuscript Submitted for Publication*.

Higgins, E. T. (1989). Self-Discrepancy Theory: What Patterns of Self-Beliefs Cause People to Suffer? *Advances in Experimental Social Psychology*, *22*, 93–136. doi: 10.1016/s0065-2601(08)60306-8

Holtgraves, T. (2004). Social Desirability and Self-Reports: Testing Models of Socially Desirable Responding. *Personality and Social Psychology Bulletin*, *30*, 161–172. doi: 10.1177/0146167203259930

Hughes, S., & Barnes-Holmes, D. (2013). A Functional Approach to the Study of Implicit Cognition: The IRAP and the REC model. *Advances in Relational Frame Theory & Contextual Behavioural Science: Research & Applications*, 97–126.

Hughes, S., Barnes-Holmes, D., & Vahey, N. (2012). Holding on to our functional roots when exploring new intellectual islands: A voyage through implicit cognition research. *Journal of Contextual Behavioral Science*, *1*, 17–38. doi: 10.1016/j.jcbs.2012.09.003

Juarascio, A. S., Forman, E. M. ., Timko, C. A., Herbert, J. D., Butryn, M., & Lowe, M. (2011). Implicit internalization of the thin ideal as a predictor of increases in weight, body dissatisfaction, and disordered eating. *Eating Behaviors*, *12*, 207–213. doi: 10.1016/j.eatbeh.2011.04.004

Parling, T., Cernvall, M., Stewart, I., Barnes-Holmes, D., & Ghaderi, A. (2012). Using the Implicit Relational Assessment Procedure to Compare Implicit Pro-Thin/Anti-Fat Attitudes of Patients With Anorexia Nervosa and Non-Clinical Controls. *Eating Disorders*, *20*, 127–143. doi: 10.1080/10640266.2012.654056

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*, 277–293. doi: 10.1037/0022-3514.89.3.277

Perkins, A. W., & Forehand, M. (2010). Implicit social cognition and indirect measures in consumer behavior. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 535–547). New York, NY: Guilford Press.

Probst, M., Vandereycken, W., Coppenolle, H. V., & Vanderlinden, J. (1995). The Body Attitude Test for Patients with an Eating Disorder: Psychometric Characteristics of a New Questionnaire. *Eating Disorders*, *3*, 133–144. doi: 10.1080/10640269508249156

Remue, J., De Houwer, J., Barnes-Holmes, D., Vanderhasselt, M.-A., & De Raedt, R. (2013). Self-esteem revisited: Performance on the implicit relational assessment procedure as a measure of self- versus ideal self-related cognitions in dysphoria. *Cognition & Emotion*, *27*, 1441–1449. doi: 10.1080/02699931.2013.786681

Remue, J., Hughes, S., De Houwer, J., & De Raedt, R. (2014). To Be or Want to Be: Disentangling the Role of Actual versus Ideal Self in Implicit Self-Esteem. *PLoS ONE*, *9*, e108837. doi: 10.1371/journal.pone.0108837

Roddy, S., Stewart, I., & Barnes-Holmes, D. (2010). Anti-fat, pro-slim, or both?: Using two reaction-time based measures to assess implicit attitudes to the slim and overweight. *Journal of Health Psychology*, *15*, 416–425. doi: 10.1177/1359105309350232

Roddy, S., Stewart, I., & Barnes-Holmes, D. (2011). Facial reactions reveal that slim is good but fat is not bad: Implicit and explicit measures of body-size bias: Facial electromyography and body-size bias. *European Journal of Social Psychology*, *41*, 688–694. doi: 10.1002/ejsp.839

Snowden, R. J., & Gray, N. S. (2010). Implicit social cognition in forensic settings. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 522–534). New York, NY: Guilford Press.

Spruyt, A., De Houwer, J., Tibboel, H., Verschuere, B., Crombez, G., Verbanck, P., … Noël, X. (2013). On the predictive validity of automatically activated approach/avoidance tendencies in abstaining alcohol-dependent patients. *Drug and Alcohol Dependence*, *127*, 81–86. doi: 10.1016/j.drugalcdep.2012.06.019

Stacy, A. W., & Wiers, R. W. (2010). Implicit cognition and addiction: a tool for explaining paradoxical behavior. *Annual Review of Clinical Psychology*, *6*, 551–575. doi: 10.1146/annurev.clinpsy.121208.131444

Stice, E. (2001). A prospective test of the dual-pathway model of bulimic pathology: Mediating effects of dieting and negative affect. *Journal of Abnormal Psychology*, *110*, 124–135. doi: 10.1037/0021-843X.110.1.124

Strauman, T. J., Vookles, J., Berenstein, V., Chaiken, S., & Higgins, E. T. (1991). Self-discrepancies and vulnerability to body dissatisfaction and disordered eating. *Journal of Personality and Social Psychology*, *61*, 946–956. doi: 10.1037/0022-3514.61.6.946

Teachman, B. A., Cody, M. W., & Clerkin, E. M. (2010). Clinical applications of implicit social cognition theories and methods. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 489–564). New York, NY: Guilford Press.

Thompson, M. A., & Gray, J. J. (1995). Development and Validation of a New Body-Image Assessment Scale. *Journal of Personality Assessment*, *64*, 258–269. doi: 10.1207/s15327752jpa6402_6

Watts, K., Cranney, J., & Gleitzman, M. (2008). Automatic evaluation of body-related images. *Body Image*, *5*, 352–364. doi: 10.1016/j.bodyim.2008.06.001

Wiers, R. W., Houben, K., Roefs, A., De Jong, Peter, & Etchison, S. (2010). Implicit cognition in health psychology: why common sense goes out the window. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 463–488). New York, NY: Guilford Press.

Wiers, R. W., Van Woerden, N., Smulders, F. T. Y., & De Jong, P. J. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology*, *111*, 648–658. doi: 10.1037/0021-843X.111.4.648

Williamson, D. A., Gleaves, D. H., Watkins, P. C., & Schlundt, D. G. (1993). Validation of self-ideal body size discrepancy as a measure of body dissatisfaction. *Journal of Psychopathology and Behavioral Assessment*, *15*, 57–68. doi: 10.1007/BF00964324

APPENDIX

Table 2. *Target stimuli of the IRAP measures and English translations*

Thinness	Thickness	Thinness (English)	Thickness (English)
dun	vet	thin	fat
mager	dik	lean	thick
fijngebouwd	zwaarlijvig	fine boned	obese
smal	gezet	narrow	squat
tenger	volslank	slender	curvaceous
slank	mollig	slim	chubby

CHAPTER 4

THE RELATIONAL RESPONDING TASK: TOWARD A NEW IMPLICIT MEASURE OF BELIEFS [1]

ABSTRACT

We introduce the Relational Responding Task (RRT) as a tool for capturing beliefs at the implicit level. Flemish participants were asked to respond *as if* they believed that Flemish people are more intelligent than immigrants (e.g., respond "true" to the statement "Flemish people are wiser than immigrants") or to respond *as if* they believed that immigrants are more intelligent than Flemish people (e.g., respond "true" to the statement "Flemish people are dumber than immigrants"). The difference in performance between these two tasks correlated with ratings of the extent to which participants explicitly endorsed the belief that Flemish people are more intelligent than immigrants and with questionnaire measures of subtle and blatant racism. The current study provides a first step towards validating RRT effects as a viable measure of implicit beliefs.

[1] Based on De Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). *The Relational Responding Task: Toward a New Implicit Measure of Beliefs.* Frontiers in Psychology

INTRODUCTION

Implicit measures have become an important part of the psychologist's toolbox (see Gawronski & De Houwer, 2014, and Nosek, Hawkins, & Frazier, 2011, for recent reviews). Although opinions differ on what it means to say that a measure is implicit, we favor the view that implicit measures are measurement outcomes that capture the to-be-measured construct (e.g., attitudes, stereotypes, evaluation) under conditions of automaticity (e.g., even when participants have little time, are engaged in multiple tasks, are not aware of what is being measured, or do not have the intention to express the construct that is being measured; see De Houwer & Moors, 2012; De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). In most implicit measurement tasks, participants are asked to respond as quickly as possible to stimuli that appear on a computer screen. For instance, in an Implicit Association Test (IAT) designed to assess implicit self-esteem, positive words, negative words, stimuli related to the self (e.g., the first name of the participant), and stimuli related to other people (e.g., the first name of another participant) appear one by one on a computer screen. In a first critical block of trials, participants are asked to press one key as quickly as possible whenever they see a positive word or a self-related stimulus and a second key whenever they see a negative word or an other-related stimulus. In a second critical block, positive words and other-related stimuli are assigned to the first key whereas negative words and self-related items are assigned to the second key. Implicit self-esteem is inferred from the difference in performance in the two critical blocks. For instance, participants who perform better in the first relative to the second block are assumed to have more positive self-esteem than participants who perform better in the second relative to the first block (e.g., Greenwald & Farnham, 2000).

A core property of most implicit measures is that they were designed to capture associations between concepts while ignoring the way in which those concepts are related (see Hughes, Barnes-Holmes, & Vahey, 2012, for a detailed overview). For instance, the propositional beliefs "I am good" and "I want to be good" both involve a relation between "I" and "good" but differ with regard to the type of relation. Even though these beliefs are fundamentally different and might coincide with entirely different behaviors, an IAT cannot differentiate between them. The Implicit Relational Assessment Procedure (IRAP), on the other hand, does allow one to differentiate at the implicit level between beliefs that differ only with regard to the relational component (for an overview, see Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2011).[2]

Consider a version of the IRAP designed to capture the belief “I am good”. On each trial of this task, either “I am” or “I am not” appears at the top of a computer screen along with a positive or negative adjective in the middle of the screen (e.g., “good” or “bad”). In this way, the task is comprised of four different types of trials (*I am + positive; I am + negative; I am not + positive; I am not +negative*). As with the IAT, there are two different types of critical blocks. In one block of trials, participants are trained to select the response “correct” on trials that are in line with the belief “I am good” (i.e., *I am + positive*; *I am not + negative*) and to respond “false” on trials that contradict this belief (i.e., *I am not + positive* and *I am + negative*). In the second type of blocks, responses need to be in line with the belief “I am bad” (i.e., respond “correct” on *I am not + positive* and *I am + negative* trials and respond “false” on *I am + positive* and *I am not + negative* trials). The better someone performs on the first type of blocks relative to the second type of blocks, the stronger that person is assumed to hold the belief “I am good” (relative to “I am bad”).[3]

Importantly, this version of the IRAP can easily be redesigned to capture the belief “I want to be good”. The only necessary change is to replace the stimuli “I am” and “I am not” with the stimuli “I want to be” and “I do not want to be”, respectively. Once this change is in place, participants can be trained to select the response “correct” on *I want*

[2] Research on implicit measures has been conducted both within the cognitive and the functional tradition in psychology (Hughes, Barnes-Holmes, & De Houwer, 2011). The term “implicit belief” could also be used either at the mental or at the functional level of explanation. At the mental level, it refers to an automatically constructed or activated propositional representation, that is, an informational unit in memory that contains information about how concepts are related and that mediates the impact of the environment on behavior by means of automatically operating mental processes (see De Houwer, 2009, 2014). At the functional level, implicit beliefs can be defined as specific patterns of (arbitrarily applicable) relational responding, that is, behavior that is under the control of the relation between events. Given the assumption that (arbitrarily applicable) relational responding is mediated by propositional representations (Hughes et al., 2011), measures such as IRAP scores can be conceptualized both as indices of relational responding (when adopting a functional perspective) and as indices of propositional representations (when adopting a mental perspective).

[3] It is also possible to calculate an IRAP score for each trial type separately, which can be used an index for four different beliefs). In the foregoing example, these trial types would be: “I am good”, “I am bad”, “I am not good”, and “I am not bad” (see Hughes et al., 2011).

to be + positive and *I do not want to be + negative* trials and to select the response "false" on *I want to be + negative* and *I do not want to be + positive* trials. During a second type of blocks, they are required to select "correct" on *I want to be + negative* and *I do not want to be + positive* trials and to select "false" for *I want to be + positive* and *I do not want to be + negative* trials. Comparing performance across these different types of blocks provides an index of the belief "I want to be good" (relative to "I want to be bad").

The significance of distinguishing between different beliefs at the implicit level was recently demonstrated in a study by Remue, De Houwer, Barnes-Holmes, Vanderhasselt, and De Raedt (2013; see also Remue, Hughes, De Houwer, & De Raedt, 2014). Their research was triggered by the observation that implicit self-esteem as indexed by IAT scores is as positive in dysphoric students and acutely depressed patients as it is in non-dysphoric students or non-depressed control participants (e.g., De Raedt, Schacht, Franck, & De Houwer, 2006). Remue and colleagues tested the idea that standard implicit measures of self-esteem might reflect different types of beliefs in different participants. For instance, in non-dysphoric students, positive self-esteem IAT scores might reflect the belief "I am good" (i.e., actual self-esteem) whereas the same score might reflect the belief "I want to be good" (i.e., ideal self-esteem) in dysphoric students. In line with this hypothesis, they observed that dysphoric students had lower scores on an IRAP designed to capture the belief "I am good" than on an IRAP designed to capture the belief "I want to be good" whereas non-dysphoric students showed the reversed pattern.

The IRAP has been used successfully in a wide variety of contexts (see Barnes-Holmes et al., 2011; and Hughes & Barnes-Holmes, 2013, for reviews). Still it would be good to have alternative measures of implicit beliefs, if only to assess and control for method specific variance. Moreover, the IRAP is often quite difficult for participants to complete. Across many studies, a substantial part of the (university student) sample fails to complete an IRAP (e.g., more than 20% of the students in the study of Remue et al., 2013 see Hughes & Barnes-Holmes, 2013, Table 1, for an overview). The difficulty of the IRAP originates - at least in part - from the fact that the assignment of keyboard keys to the "correct" and "false" response option typically varies from trial to trial (e.g., press d for “correct” and k for “false” on trial n but press d for “false” and k for “correct” on trial n+1). The assignment of responses to different keys is varied in order to avoid a potential confound between the function of the response (indicate “correct” or “false”) and its physical location (e.g., the ‘d’ or ‘k’ keys). This confound could alter the encoding of the responses by allowing participants to select a response on the basis of its spatial properties (e.g., “if *I am + positive*, then press key d”) rather than its meaning (e.g., “*if I am + positive*, then press correct”; but see Campbell, Barnes-Holmes, Barnes-Holmes, & Stewart, 2011, for data showing that keeping the assignment of responses fixed does not necessarily reduce IRAP effect sizes). Although it might be possible to avoid a high drop-out rate by increasing prior exposure to the task or changing properties of the procedure (see Hughes & Barnes-Holmes, 2013; and Vahey, Boles, & Barnes-Holmes, 2010, for a discussion), the difficulty in completing an IRAP does constrain its utility, especially when using it via online applications (in which case extensive training is often not feasible) or with certain populations (e.g., clinical patients with limited attentional or intellectual capacities).

In the present paper, we offer the Relational Responding Task (RRT) as a novel measurement procedure that, like the IRAP, is designed to capture specific beliefs at the implicit level. It retains an essential ingredient of the IRAP – namely – the requirement for participants to respond in-line with specific beliefs (e.g., “I am good” or “I want to be good”). Unlike most instantiations of the IRAP, the RRT involves the presentation of full statements in the middle of the computer screen (e.g., “I like myself”) and participants are explicitly instructed to act “as if” they agree with certain statements and disagree with others.[4] For instance, in a first block, participants might be asked to respond as if they believe that they are good by selecting “true” when presented with statements that imply positive views of the self (e.g., “I like myself”) and by selecting “not true” when presented with statements that imply negative views of the self (e.g., “I dislike myself”). In a second block, they would respond as if they believe that they are bad by selecting “true” when presented with negative self-related statements and “false” when presented with positive-self related statements. The

[4] In many instantiations of the IRAP, participants are trained to respond in certain ways via error feedback rather than given explicit instructions about the stimulus-response assignments.

difference in performance between these two types of blocks is assumed to provide a measure of the extent to which participants believe that they are good.

In addition to differences in the type of stimuli used and the nature of the instructions, the IRAP and RRT also differ in their structural properties. In the RRT, the physical properties of the correct and false response remain constant throughout the task. On all trials, responding “true” is realized by pressing a first key whereas responding “not true” is done by pressing a second key. Recoding of the responses (e.g., in terms of physical location) is discouraged, however, by including inducer trials on which stimuli are presented that refer to the concepts “true” or “not true” (e.g., words such as “correct” and “wrong”). Participants are instructed to press one key for inducer stimuli that refer to “true” and to press a second key for inducer stimuli that refer to “false”. Similar inducer trials have been used successfully in other implicit measurement procedures as a means to discourage recoding of the responses (e.g., Extrinsic Affective Simon Task; De Houwer, 2003).

Interestingly, the inclusion of inducer trials results in the RRT having a task structure that closely resembles the IAT. More specifically, in both the RRT and the IAT, four categories of stimuli are assigned to two responses in a way that varies across blocks. As a result, it is likely that the difficulty of performing an RRT will be similar to the difficulty of performing an IAT. Given that (a) attrition rates in IAT studies are typically low, that (b) IATs have already been successfully administrated via the world wide web on a large scale, and that (c) IATs have been used successfully with a variety of populations (e.g., Nosek, Greenwald, & Banaji, 2007, for a review), we believe that the convenience of use of the RRT will be high, and in several respects, superior to that of the IRAP (at least as currently instantiated).

Despite their structural similarity, it is important to realize that the IAT is fundamentally different from the RRT (and IRAP) in several ways. Most importantly, a typical IAT does not require participants to relate the different categories that are used in the IAT. In a self-esteem IAT, for example, participants can select the correct response merely by identifying the individual category that the presented stimulus is a member of (e.g., the word “me” is a member of the category of stimuli labeled “self”).

In the RRT (and the IRAP), on the other hand, participants can only select the correct response based on the way that different stimuli are related to one another. For instance, in an RRT designed to assess the belief "I am good", a correct response does not simply depend on the presence of specific stimuli (e.g., "I", "am", or "good") but on the relation *between* different stimuli (e.g., the statement "I am good"). Moreover, the instruction to "act as if" statements are true or false encourages participants to relate the presented statement with the categories "true" and "not true". That is, participants need to respond to a statement as being either true or false based on the rule that is specified for each block of trials. In sum, the RRT requires participants to respond in a complex relational manner (hence the name *relational responding* task). It is precisely this feature of the RRT that endows it with the potential to capture individual beliefs that differ with regard to the relational component (e.g., "I am good" versus "I want to be good").

In the current paper, we set out to validate empirically the RRT effect as an implicit measure of beliefs. Rather than immediately testing the potential of the RRT to capture differences between beliefs that vary only in their relational component (e.g., by comparing an RRT designed to capture the belief "I am good" with an RRT that captures the belief "I want to be good"), we first conducted a study that focused on a single belief and examined whether an RRT measure of that belief correlates with criterion variables (see De Houwer et al., 2009, for more information on how to validate implicit measures). More specifically, we used an RRT that was designed to capture the extent to which Flemish participants hold the belief that Flemish people are more intelligent than immigrants. The inducer stimuli were synonyms of the concepts "true" or "false" that had to be categorized as either "true" by pressing one key or as "false" by pressing another key. Target trials involved the presentation of statements that were either in line with the belief that Flemish people are more intelligent that immigrants (e.g., "Flemish people are smarter than immigrants", "Immigrants are dumber than Flemish people") or in line with the belief that immigrants are more intelligent than Flemish people (e.g. "Flemish people are dumber than immigrants", "Immigrants are smarter than Flemish people").

In addition to practice blocks that contained either inducer trials or target trials only, the RRT also included two types of test blocks that contained both inducer and target trials. During the first type of test block, participants were asked to respond to the target statements as if they believed that Flemish people are more intelligent than immigrants (i.e., to respond “true” to statements such as “Flemish people are smarter than immigrants” and to respond “not true” to statements such as “Flemish people are dumber than immigrants”; pro-Flemish block). In the second type of test block, participants were asked to respond as if they believed that immigrants are more intelligent than Flemish people (i.e., to respond “not true” to statements like “Flemish people are smarter than immigrants” and to respond “true” to statements like “Flemish people are dumber than immigrants”; pro-immigrant block). The difference in performance across these two types of blocks is assumed to assess the belief that Flemish people are more intelligent than immigrants.

In order to validate the RRT scores, we also asked participants to express on a rating scale the extent to which they believe that Flemish people are more or less intelligent than immigrants. Because implicit and explicit measures of the same construct typically converge to some extent (Nosek, 2007), we expected that the RRT and rating measure would be correlated. Participants also completed questionnaires designed to capture subtle, modern, and blatant forms of racism. Blatant racism is considered to be a hot and direct form of racial prejudice, often referring to beliefs about the (genetic) inferiority of the racial out-group. It is typically associated with a strong opposition to any intimate contact with the out-group. Subtle racism, on the other hand, is a more cool and covert expression of racial prejudice, including the idea that out-groups are poorly adapted to the ingroup’s traditional values, an exaggeration of cultural differences between the in-group and the out-group, and the absence of positive emotions (rather than the presence of negative emotions) toward the out-group (see Pettigrew & Meertens, 1995). Similarly, modern racism also refers to a more covert form of prejudice (as opposed to the blatant, old-fashioned forms) but is distinct in that it taps into beliefs that racial discrimination is no longer a problem (or even does not exist anymore) and that out-groups (i.e., Black people in the original work) have become too demanding and are pushing for unfair advantages (McConahay, 1986). If a

Flemish person believes that Flemish people are more intelligent than immigrants, this belief could be regarded as an instance of racial prejudice, especially blatant/old-fashioned racism (e.g., Bobo & Kluegel, 1993; McConahay, 1986). We therefore expected that RRT scores would also correlate with scores on these racism questionnaires, in particular blatant racism.

Finally, please note that we did not make predictions about the direction or magnitude of the overall RRT score. First, the overall RRT score in our study would probably be biased as the result of order effects. We fixed the order of the blocks (i.e., all participants started with the pro-Flemish block) because we were interested primarily in interindividual differences in RRT scores, that is, we wanted to correlate RRT scores with other measures. Counterbalancing block order is known to increase error variance (i.e., differences in scores between participants might reflect not only differences in the to-be-measured attribute but also differences in block order) and thus to lower correlations (see Perugini & Banse, 2007, for a discussion and the recommendation to fix block order). Because there are few studies about the effect of block order on the validity of implicit measures (e.g., studies examining whether correlations with validity criteria are stronger when starting with an attitude-inconsistent versus an attitude-consistent block), we did not have strong reasons to select one of the two block orders but more or less randomly decided to always start with the pro-Flemish block. More importantly, fixing the order of blocks complicates the interpretation of overall scores because those scores could be influenced by the order in which the blocks are completed. For instance, practice with the stimulus-response mappings in the first critical block might slow down responding in the second critical block during which those mappings are reversed. On the other hand, performance in the second critical block might be facilitated because of practice effects (e.g., faster responding because items are more familiar and thus easier to process). When block order is fixed, it is impossible to determine to what extent the overall score reflects the reversal of stimulus-response mappings, practice effects, or the properties of the to-be-measured attribute. Hence, it is best to refrain from interpreting the overall score when block order is fixed. A second reason for refraining from an interpretation of the overall RRT score is that, even without biases due to fixed block order, it is difficult to determine the correct interpretation of the zero point on psychological measures (Blanton & Jaccard, 2006).

METHOD

Participants

Forty-nine students at Ghent University (M_{age} = 23 years, SD_{age} = 4; five men) participated in exchange for €5. Ghent University is situated in Flanders, which is the northern region of Belgium. These participants were classified as being Flemish because they indicated that Dutch (which is the dominant language in Flanders) was their native language and that both their parents were Flemish. The data of six other participants who told the experimenter that they had at least one parent of non-Belgian nationality or said that Dutch was not their (only) native language, were excluded from the analysis. All participants reported normal or corrected-to-normal vision and provided their informed consent before participating.

Materials

RRT. Ten words and twenty statements were used as stimuli. Five words were related to “true” (the Dutch words “goed”, “juist”, “correct”, “exact”, and “in orde”) and five words were related to “false” (the Dutch words “mis”, “onjuist”, “incorrect”, “verkeerd”, and “fout”). These words were presented during the inducer trials. Each of the twenty target statements (see Appendix) related *Flemish people* and *immigrants* to one another in terms of their intelligence level, using five synonyms for intelligent (e.g., smarter) and five synonyms to denote a lack of intelligence (e.g., dumber). Ten statements implied that Flemish people are more intelligent than immigrants (e.g., “Flemish people are smarter than immigrants”, “Immigrants are less clever than Flemish people”) while another ten statements implied that immigrants are more intelligent than Flemish people (e.g., “Immigrants are wiser than Flemish people”, “Flemish people are less wise than immigrants”). All statements were presented in bold Verdana font, size 28. Inducer stimuli were presented in orange, whereas target statements were presented in blue throughout the task. The experiment was conducted using a 17 inch LCD screen (60 Hz, 1440 × 900 pixels). The RRT program was written in Affect 4.0 (Spruyt, Clarysse, Vansteenwegen, Baeyens, & Hermans, 2010).

Questionnaires. Participants completed a 12-item subtle racism scale and an 8-item blatant racism scale (Pettigrew & Meertens, 1995; adapted by Van Hiel & Mervielde, 2005), along with a 10-item modern racism scale (McConahay, 1986; translated by Dhont, Cornelis, & Van Hiel, 2010). Virtually all questions referred to immigrants in general, with the exception of a few questions that referred to Turks and Moroccans, the two largest groups of immigrants in Belgium from outside of the European Union. Also, a few items stated that immigrants differ in their religious beliefs or culture from Belgians, implying that immigrants do not include foreigners from other western European countries. Examples are "Immigrants should be wise enough not to impose themselves at places where they know beforehand that they would be discriminated" (subtle racism scale), “Discrimination against immigrants no longer is problem in Belgium" (modern racism scale), and “Immigrants are a threat to the employment of Belgians" (blatant racism). Note that some questions referred to Belgian people as the in-group whereas the RRT referred to Flemish people as the in-group. If anything, this difference would reduce the magnitude of the correlations between the racisms measures and the RRT and thus work against our hypotheses. However, we did not expect a strong negative impact of this divergence because our participants are both Belgian and Flemish. Hence, both labels describe the in-group correctly as opposed to the out-group of immigrants.

All questions were rated on 7-point Likert scales ranging from 1 (‘completely disagree’) to 7 (‘completely agree’). In addition, participants were asked to provide their opinion on the relative intelligence of Flemish people and immigrants using a scale that ranged from -10 (immigrants are more intelligent than Flemish people) to +10 (Flemish people are more intelligent than immigrants), with 0 indicating that the two groups do not differ in their intelligence.

Procedure

Participants were tested individually in a dimly lit room. All participants first completed the RRT. They were instructed to categorize words and statements presented on the screen as either “true” or “not true” by pressing the left or right control-keys of the keyboard, respectively. In the first block (40 trials), each of the 10 inducer words was presented 4 times as practice. Participants were asked to categorize the words as synonyms of “true” (press right control key) and “not true” (press left control key). In the second block of 40 trials, the 20 target statements were presented twice. Participants were asked to respond to the statements in line with the rule that Flemish people are more intelligent than immigrants (and immigrants are therefore less intelligent than Flemish people). The third block consisted of two consecutive repetitions of 40 trials, in which the 10 inducer stimuli were presented twice and the 20 target statements were presented once, leading to a total of 80 trials. Participants were asked to respond in accordance with the rules practiced in the two preceding blocks. Block 4 was identical to the second block except for a reversal of the rule for responding. That is, participants were asked to respond as if immigrants are more intelligent than Flemish (and Flemish are therefore less intelligent than immigrants). The fifth and final block was identical to the third, but participants were asked to respond to target statements in accordance with the response rule learned in block four.

The order of the trials was random except for the restriction that the same statement or word could not be presented on two consecutive trials. Each trial started with the presentation of a word or statement in the middle of the screen. It remained there until a response was registered. Incorrect responses were followed by the presentation of a red cross which remained on screen until participants gave the appropriate response. The following trial started 750 ms after a correct response was emitted. Once they completed the RRT, participants rated the target belief after which they completed the three questionnaires (subtle, modern, and blatant racism scale, in that order). They then were debriefed and paid.

RESULTS

RRT data

Given that inducer trials were included merely to prevent response recoding, only data from the target trials of the mixed blocks were analyzed. Analyses that included also the data of the inducer trials led to the same conclusions. The data from two participants were excluded from the analyses because their percentage of errors was more than 2.5 standard deviations above the mean percentage of errors of the total group. Reaction times were defined as the time in milliseconds between the onset of presentation of the statement and the registration of the correct response. These reaction times were transformed into D_{RRT} scores using the same improved D-algorithm (D1) that Greenwald, Nosek, and Banaji (2003) developed for the IAT. The D1 score was chosen because we recorded reaction times until the correct response was emitted, thus removing the need for algorithms that add extra penalty time to reaction times on trials with an incorrect response. Analyses using the D4 (also known as D600) scoring algorithm, led to similar conclusions. D_{RRT} effects were scored so that positive values indicated faster responses in the pro-Flemish block (i.e., act as if Flemish people are more intelligent than immigrants) relative the pro-immigrant block of the RRT (i.e., act as if immigrants are more intelligent than Flemish people) while negative values indicated the opposite. The D_{RRT} score ranged from -0.79 to 0.56, with a mean D_{RRT} score of M = -0.01 (SD = 0.34). The mean D_{RRT} score did not differ significantly from zero, $t < 1$. During the mixed blocks, participants on average took 1551 ms (SD = 373.93) to respond to the target statements and 672 ms (SD = 88.58) to respond to the inducer words. In these blocks, they made on average 11.1 % of errors (SD = 0.06) on trials with target statements and 5.3% of errors (SD = 0.04) on trials with inducer words.

Correlation analyses

Table 1 provides an overview of all pairwise correlations between the D_{RRT} score, the explicit rating, and the three questionnaires. Most importantly, the D_{RRT} score correlated significantly with the explicit judgments of the relative intelligence of Flemish

Table 1. *Correlations between measures.*

		1	2	3	4	5
1	D_{RRT}	-	0.43*	0.36*	0.20	0.34*
2	Rating		-	0.40*	0.31*	0.70*
3	Subtle Racism			-	0.74*	0.80*
4	Modern Racism				-	0.67*
5	Blatant Racism					-

* $p < .05$

and immigrants, $r = 0.43$, $t(44) = 3.17$, $p = .003$, with the scores on the subtle racism scale, $r = 0.36$, $t(45) = 2.59$, $p = .013$, and with the scores on the blatant racism scale, $r = 0.34$, $t(45) = 2.39$, $p = .020$. The correlations between the D_{RRT} score and those obtained from the modern racism scale did not reach conventional levels of significance, $r = .20$, $t(45) = 1.39$, $p = .171$.

Odd-even split half reliability of the RRT score, using Spearman-Brown correction, was $Rsb = 0.64$. For the questionnaire of subtle, modern, and blatant racism, Chronbach's alpha was .85, .83, and .86, respectively. The mean value on those questionnaires was 4.26 ($SD = .86$), 3.55 ($SD = .87$), and 2.61 ($SD = 1.05$), respectively. Finally, the mean score on the explicit rating was 1.28 ($SD = 1.05$) which differed from zero, $t(45) = 4.41$, $p < .001$.

DISCUSSION

The current study sought to provide a first step toward the validation of the RRT effect as an implicit measure of beliefs. In line with the hypothesis that the RRT scores capture the (prejudiced) belief that Flemish people are more intelligent than immigrants, we observed a significant correlation between RRT scores and explicit ratings of how much participants endorsed this belief. RRT scores also correlated positively with questionnaires of subtle and blatant (but not modern) racism. This pattern of correlations supports the claim that the RRT effects captured the to-be-measured belief and hence provides preliminary evidence for the idea that the RRT can provide a general tool for assessing beliefs. In addition, the RRT seems to be more user friendly than the IRAP. First, the attrition rate in our study was low. All participants successfully completed the RRT. We did discard the data of two participants because their percentage of errors was more than 2.5 standard deviations above the mean percentage of errors of the total group. However, this still implies an attrition rate of less than 5% (i.e., 4.08%) whereas attrition rates of 20% or more are common in IRAP studies. Moreover, participants needed only about 10 minutes to complete the RRT whereas it often takes 20 minutes or more to complete the various stages of the IRAP.

We not only put forward the claim that RRT scores capture beliefs but also that they provide an *implicit* measure of those beliefs. It has been argued that a measure qualifies as implicit if it functions as a valid measure even under conditions of automaticity (De Houwer & Moors, 2012; De Houwer et al., 2009). Based on the structural features of the RRT, one could argue that RRT scores provide an implicit measure in the sense that they capture beliefs even though participants are asked to respond quickly. Speed is of course a continuous variable, which implies that measures differ in the degree to which they are implicit along this criterion. In our study, for example, participants required an average of 1551 ms to respond to the target statements and 672 ms to respond to inducer words. Although 1551 ms might seem long relative to the reaction times observed in most other implicit measurement tasks, the target statements used in our study are much more complex than the single words or pictures that are usually presented in other implicit measurement tasks. Hence, it seems likely that participants did respond very quickly once they processed the meaning of the statements. Consistent with this conclusion, reaction times on inducer trials (during which participants reacted to individual words) were much shorter and comparable to those seen in other implicit measurement tasks. Regardless of how RRT scores relate to other implicit measures in this regard, it seems safe to argue that RRT scores are more implicit in terms of speed than scores obtained from traditional self-report procedures (e.g., ratings, questionnaires) that allow for ample time to reflect and respond to questions.

Are RRT measures implicit in the sense of unintentional? Unlike explicit measures such as the rating measures used in our study, participants are not asked to express their beliefs during an RRT. Instead, they are simply asked to *act as if* they endorse a certain belief. On the one hand, when asked to act in ways that contradict their beliefs, it is unlikely that participants will express those beliefs in an intentional manner because this would lead to incorrect responses. On the other hand, when asked to respond in line with beliefs that they do endorse, it is possible (but not necessarily the case) that participants do express their beliefs in an intentional manner rather than act as if they endorse those beliefs. Hence, at present, it is difficult to make strong claims about whether RRT measures are implicit in the sense of valid in the absence of the intention to express a belief. Future work could examine this issue further by determining the strategies that participants use while completing an RRT. Such work could also examine if the RRT meets other conditions of automaticity, such as controllability.

What would it mean to say that RRT scores capture beliefs at the implicit level? One way to conceptualize this statement is in terms of the implicit endorsement of propositions. At first sight, the idea that propositions can be endorsed implicitly might seem self-contradictory in light of cognitive theories that postulate that all implicit processing relies on associative representations (e.g., Rydell & McConnell, 2006). Unlike propositions, associations do not have a truth value: an association does not imply a statement about the world and is therefore neither true nor untrue. Hence, associations as such cannot be endorsed (i.e., be evaluated as true).[5] Together, these two assumptions (all implicit processing relies on associations; associations cannot be endorsed) imply that endorsement can never be implicit.

More recently, however, cognitive researchers have raised the possibility that implicit processing can involve propositional representations (De Houwer, 2014; Hughes, Barnes-Holmes, & De Houwer, 2011; Hughes et al., 2012). Indeed, a rapidly growing body of evidence supports the conclusion that propositions can be both formed and activated under the various conditions of automaticity (see De Houwer, 2014, for a review). At the same time, Shidlovski, Schul, and Mayo (2014), recently advanced the concept of "implicit truth" which they define as the automatic endorsement of propositions. Across several studies, the authors showed that indices of implicit truth can be dissociated from the explicit endorsement of truth. From this

[5] The proposition that an association exists can be endorsed but the association itself cannot.

perspective, scores on the RRT (and IRAP) do have a unique and potentially crucial role to fulfill in psychological research as measures of the implicit endorsement of propositions.

One could argue, however, that the IRAP and RRT are not alone in their ability to capture beliefs at the implicit level. For instance, several variants of the IAT use categories that specify relational information. Consider the so-called personalized IAT (pIAT) in which participants classify (attribute) stimuli as instances of the categories “I like” or “I dislike” rather than the categories “good” or “bad” (Olson & Fazio, 2004; see Dewitte & De Houwer, 2008; and Yoshida, Peach, Zanna, & Spencer, 2012 for related propositionalized variants of the IAT). For instance, in a smoking pIAT (e.g., De Houwer, Custers, & De Clercq, 2006), participants see words that refer to things they like (e.g., flowers), words that refer to things they dislike (e.g., cockroaches), pictures related to smoking (e.g., a lighter), and pictures unrelated to smoking (e.g., a pencil). One could argue that, in this case, the categorization of words is no longer based on category membership as such (i.e., good, bad, or I) but rather on the basis of the relation between different concepts (“I” in combination with “like” or “dislike”).

Nevertheless, while performance in the pIAT might be relationally more complex than performance on traditional IATs, the pIAT remains fundamentally different from the RRT (and IRAP). Like other IATs, the pIAT does not require participants to relate the different components of a to-be-measured belief. For instance, in a smoking pIAT, response selection is determined on the basis of individual elements of the beliefs “I like smoking” or “I dislike smoking” (i.e., the individual categories “I like”, “I dislike”, “smoking”, or “non-smoking”). In the RRT, on the other hand, response selection is instructed in terms of the full combination of these elements (i.e., “I like smoking” or “I dislike smoking”). Moreover, as we pointed out in the introduction, in the RRT but not in IATs, participants are instructed to respond on the basis of whether a statement is to be considered as true or false in a specific block of trials.

The autobiographical IAT (aIAT; see Verschuere, Suchotzki, & Debey, 2014, for a review) is another variant of the IAT that might allow for the measurement of beliefs and seems to bear even greater similarity to the RRT at the structural level than other

IATs. In an aIAT, participants encounter generic statements that are true (e.g., “I am a human being”) or false (e.g., “I am a cow”) for all participants. They are required to classify these statements as either “true” or “false” by pressing one of two keys. Participants also see statements that describe a particular event (Event A) that the participant has witnessed (autobiographical items; e.g., statements related to a crime) or a second event (Event B) that the participant has not witnessed (control items). In one of the critical test blocks, participants press the same key for true items and items referring to Event A and a second key for false items and items referring to Event B. In the second critical test block, the stimulus-response assignments are reversed (press the first key for true items and Event B items; press the second key for false items and Event A items). As is the case for the pIAT, the aIAT is relationally more complex than traditional IATs. Most importantly, in the aIAT, participants respond to statements (i.e., combinations of items) rather than individual elements of a statement. Nevertheless, unlike the RRT, the aIAT does not require participants to relate all elements of the to-be-measured belief (e.g., “Event A is true”) because responses can be selected on the basis of the individual elements of that belief (e.g., “Event A” or “true”). Therefore, like all other variants of the IAT, the aIAT does not require participant to respond on the basis of whether a statement is to be considered as true or false.

The fact that the RRT is fundamentally different from both the pIAT and aIAT does not, however, allow for the conclusion that only the RRT can be used to capture beliefs at the implicit level. In fact, we believe that scores on the pIAT, aIAT, and even traditional IATs reflect beliefs (also see De Houwer, 2014). Although IATs do not require participants to relate all elements of a to-be-measured belief, participants might still relate those elements to one another, either explicitly or implicitly. Take the self-esteem IAT that we mentioned in the introduction. Rather than responding to self-related items merely on the basis of the fact that the item refers to oneself, participants could recode the task in such a way that they respond to self-related items as positive items (i.e., based on the fact that self-related items such as “I” have a positive valence for the participant). This would imply that self-related items are responded to in terms of how they relate to positive and negative valence, which could reflect beliefs about those items (e.g., “I like myself”). Importantly, the way in which items are responded to can vary in implicit, unintentional ways. For instance, participants might sometimes erroneously classify self-related items on the basis of their valence even when they do not have the intention to do so. In fact, one of the most successful accounts of the IAT implies that IAT effects are due to such instances of task misapplication (i.e., Klauer & Mierke, 2005). Hence, IAT effects might well reflect the implicit (in the sense of fast and unintentional) endorsement of propositions.

Because traditional IATs do not incorporate any information about the relation between the different categories, different participants might (implicitly or explicitly) relate the categories of the IAT in different ways. For instance, whereas non-dysphoric students might relate items that refer to the self and positive items in terms of liking (i.e., respond to items in terms of whether they refer to something you like), dysphoric students might relate those items in terms of aspirations (i.e., respond to items in terms of whether they are something you want). Although this idea leaves open the question of why dysphoric and non-dysphoric participants differ in this manner, it is in line with the fact that self-esteem IAT scores do not differ between those participants even though all other evidence leads to the conclusion that dysphoric participants do have lower self-esteem (De Raedt et al., 2006). It is also in line with the IRAP findings of Remue and colleagues (Remue et al., 2013, 2014) which suggest that dysphoric and non-dysphoric students do indeed differ in their implicit endorsement of the propositions “I am good” and “I want to be good”. From this perspective, pIATs and aIATs differ from traditional IATs in that the former are more likely to guide the way in which participants relate the different categories by specifying the precise way in which those stimuli should be related (e.g., “I like” or “I dislike”).

In short, it seems reasonable to assume that scores on IATs (and other implicit measures; see De Houwer, 2014) reflect beliefs at the implicit level. Still, the RRT is likely to have advantages over other implicit measures when it comes to capturing beliefs. Most importantly, because the RRT requires participants to relate the elements of beliefs in a highly specific manner, the RRT offers more control over the way that participants relate stimuli and thus over the type of beliefs that scores on the RRT capture. Of course, whether the RRT (or IRAP) outperforms the IAT as an implicit measure of beliefs is an empirical issue.

Finally, a reviewer alerted us to the fact that the RRT bears some resemblance to one particular variant of the IRAP in which participants are cued to lie or respond truthfully (Levin, Hayes, & Waltz, 2010). When the word LIE is presented on the screen, participants select response “similar/yes” if they consider the two concepts on the screen to be dissimilar (e.g., “addict” and “good”) and to select the response “dissimilar/no” if they consider the concepts to be similar (e.g., “addict” and “bad”). On trials with the word TRUTH, “similar/yes” has to be selected for similar pairs and “dissimilar/no” for dissimilar pairs. Unlike to what is the case in the RRT, in this variant of the IRAP, the experimenter does not determine which responses are correct or incorrect. Instead, it is the participant who first has to decide what is true for him or her, after which he or she can select the response in line with the cue on that trial (TRUTH or LIE). Hence, this version of the IRAP seems to tap into how well people can intentionally lie about their beliefs (relative to telling the truth). In the RRT, on the other hand, participants are instructed by the experimenter to respond in a particular manner that might be either consistent or inconsistent with the beliefs of the participant. The task does not require participants to decide what their beliefs are. Which responses qualify as correct depends entirely on the experimenter. As such, RRT effects do not capture the ability to lie in an intentional manner (i.e., to retrieve and then conceal the truth) but the ability to act in ways that are consistent or inconsistent with a belief one might have. Although it needs to be settled empirically which of the two tasks is more valid and useful, this variant of the IRAP does differ from the RRT in non-trivial ways. Moreover, as is the case in most instantiations of the IRAP, attrition rates in this variant were very high (25%), probably because of the trial-by-trial changes in the location of the response options and the nature of response cue. Because we observed much a lower attrition rate in our RRT study, it does seem to be the case that the RRT is more user friendly than the variant of the IRAP that was introduced by Levin et al. (2010).

To conclude, the present paper introduces the RRT as a tool for obtaining an implicit measure of beliefs. We reported a first study in which the RRT was used to measure the extent to which Flemish participants hold the belief that Flemish people are more intelligent than immigrants. As expected, scores on the RRT correlated with ratings of how much participants explicitly endorse this belief. Also correlations with questionnaires of racism were observed. Although these findings provide initial evidence for the claim that the RRT provides a way to capture beliefs, more research is clearly needed, especially research in which RRT scores are related with real-life behavior. Indeed, it might well be that RRT effects allow one to predict unique variance in behavior that cannot be predicted on the basis of explicit measures of beliefs. Such an added value is likely to arise in situations where implicitly endorsed propositions differ from explicitly endorsed propositions (see Shidlovski et al., 2014). Future work will also need to substantiate the claim that RRT scores qualify as an *implicit* measure of beliefs and to compare the RRT with other potential implicit measures of beliefs. Moreover, it would be good to directly compare RRT measures with other measures in terms of validity, utility, and employability. By providing the first step in the validation process, however, the present paper highlights the potential of the RRT and sets the stage for such work.

REFERENCES

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2011). A sketch of the Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *The Psychological Record*, *60*, 527–542.

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist*, *61*, 27–41. doi: 10.1037/0003-066X.61.1.27

Bobo, L., & Kluegel, J. R. (1993). Opposition to race-targeting: Self-interest, stratification ideology, or racial attitudes? *American Sociological Review*, *58*, 443–464. doi: 10.2307/2096070

Campbell, C., Barnes-Holmes, Y., Barnes-Holmes, D., & Stewart, I. (2011). Exploring Screen Presentations in the Implicit Relational Assessment Procedure (IRAP). *International Journal of Psychology and Psychological Therapy*, *11*, 377–388.

De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology*, *50*, 77–85. doi: 10.1026//1618-3169.50.2.77

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*, 1–20. doi: 10.3758/LB.37.1.1

De Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass*, *8*, 342–353. doi: 10.1111/spc3.12111

De Houwer, J., Custers, R., & De Clercq, A. (2006). Do smokers have a negative implicit attitude toward smoking? *Cognition & Emotion*, *20*, 1274–1284. doi: 10.1080/02699930500484506

De Houwer, J., & Moors, A. (2012). How to define and examine implicit processes. In R. Proctor & J. Capaldi (Eds.), *Psychology of science: Implicit and explicit processes* (pp. 183–198). Oxford: Oxford University Press.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368. doi: 10.1037/a0014211

De Raedt, R., Schacht, R., Franck, E., & De Houwer, J. (2006). Self-esteem and depression revisited: Implicit positive self-esteem in depressed patients?

Behaviour Research and Therapy, *44*, 1017–1028. doi: 10.1016/j.brat.2005.08.003

Dewitte, M., & De Houwer, J. (2008). Proximity and distance goals in adult attachment. *European Journal of Personality*, *22*, 675–694. doi: 10.1002/per.696

Dhont, K., Cornelis, I., & Van Hiel, A. (2010). Interracial public–police contact: Relationships with police officers' racial and work-related attitudes and behavior. *International Journal of Intercultural Relations*, *34*, 551–560. doi: 10.1016/j.ijintrel.2010.07.004

Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (Vol. 2, pp. 283–310). New York, NY: Cambridge University Press.

Greenwald, A. G., & Farnham, S. D. (2000). Using the implicit association test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, *79*, 1022–1038. doi: 10.1037/0022-3514.79.6.1022

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216. doi: 10.1037/0022-3514.85.2.197

Hughes, S., & Barnes-Holmes, D. (2013). A Functional Approach to the Study of Implicit Cognition: The IRAP and the REC model. *Advances in Relational Frame Theory & Contextual Behavioural Science: Research & Applications*, 97–126.

Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorizing in implicit attitude research: Propositional and behavioral alternatives. *The Psychological Record*, *61*, 464–498.

Hughes, S., Barnes-Holmes, D., & Vahey, N. (2012). Holding on to our functional roots when exploring new intellectual islands: A voyage through implicit cognition research. *Journal of Contextual Behavioral Science*, *1*, 17–38. doi: 10.1016/j.jcbs.2012.09.003

Klauer, K. C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task-set switching account of the Implicit Association Test effect. *Personality and Social Psychology Bulletin*, *31*, 208–217. doi: 10.1177/0146167204271416

Levin, M. E., Hayes, S. C., & Waltz, T. (2010). Creating an implicit measure of cognition more suited to applied research: A test of the Mixed Trial—Implicit Relational Assessment Procedure (MT-IRAP). *International Journal of Behavioral Consultation and Therapy*, *6*, 245–262. doi: 10.1037/h0100911

McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination and racism* (pp. 91–126). New York, NY: Academic.

Nosek, B. A. (2007). Implicit–explicit relations. *Current Directions in Psychological Science*, *16*, 65–69. doi: 10.1111/j.1467-8721.2007.00477.x

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes* (pp. 265–292). New York, NY: Psychology Press.

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: from measures to mechanisms. *Trends in Cognitive Sciences*, *15*, 152–159. doi: 10.1016/j.tics.2011.01.005

Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extrapersonal associations on the Implicit Association Test: personalizing the IAT. *Journal of Personality and Social Psychology*, *86*, 653–667. doi: 10.1037/0022-3514.86.5.653

Perugini, M., & Banse, R. (2007). Personality, implicit self-concept and automaticity. *European Journal of Personality*, *21*, 257–261. doi: 10.1002/per.637

Pettigrew, T. F., & Meertens, R. W. (1995). Subtle and blatant prejudice in Western Europe. *European Journal of Social Psychology*, *25*, 57–75. doi: 10.1002/ejsp.2420250106

Remue, J., De Houwer, J., Barnes-Holmes, D., Vanderhasselt, M.-A., & De Raedt, R. (2013). Self-esteem revisited: Performance on the implicit relational assessment procedure as a measure of self- versus ideal self-related cognitions in dysphoria. *Cognition & Emotion*, *27*, 1441–1449. doi: 10.1080/02699931.2013.786681

Remue, J., Hughes, S., De Houwer, J., & De Raedt, R. (2014). To Be or Want to Be: Disentangling the Role of Actual versus Ideal Self in Implicit Self-Esteem. *PLoS ONE*, *9*, e108837. doi: 10.1371/journal.pone.0108837

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: a systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*, 995–1008. doi: 10.1037/0022-3514.91.6.995

Shidlovski, D., Schul, Y., & Mayo, R. (2014). If I imagine it, then it happened: The Implicit Truth Value of imaginary representations. *Cognition*, *133*, 517–529. doi: 10.1016/j.cognition.2014.08.005

Spruyt, A., Clarysse, J., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2010). Affect 4.0: A Free Software Package for Implementing Psychological and Psychophysiological Experiments. *Experimental Psychology (formerly Zeitschrift Für Experimentelle Psychologie)*, *57*, 36–45. doi: 10.1027/1618-3169/a000005

Vahey, N., Boles, S., & Barnes-Holmes, D. (2010). Measuring adolescents' smoking-related social identity preferences with the Implicit Relational Assessment Procedure (IRAP) for the first time: A starting point that explains later IRAP evolutions. *International Journal of Psychology and Psychological Therapy*, *10*, 453–474.

Van Hiel, A., & Mervielde, I. (2005). Authoritarianism and social dominance orientation: Relationships with various forms of racism. *Journal of Applied Social Psychology*, *35*, 2323–2344. doi: 10.1111/j.1559-1816.2005.tb02105.x

Verschuere, B., Suchotzki, K., & Debey, E. (2014). Detecting Deception Through Reaction Times. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Detecting Deception: Current Challenges and Cognitive Approaches* (pp. 269–291). Chichester, UK: John Wiley & Sons.

Yoshida, E., Peach, J. M., Zanna, M. P., & Spencer, S. J. (2012). Not all automatic associations are created equal: How implicit normative evaluations are distinct from implicit attitudes and uniquely predict meaningful behavior. *Journal of Experimental Social Psychology*, *48*, 694–706. doi: 10.1016/j.jesp.2011.09.013

APPENDIX

Table 2. *Target statements of the RRT and English translations.*

Target statements	English translations
Vlamingen zijn intelligenter dan allochtonen.	Flemish people are more intelligent than immigrants.
Vlamingen zijn slimmer dan allochtonen.	Flemish people are smarter than immigrants.
Vlamingen zijn verstandiger dan allochtonen.	Flemish people are wiser than immigrants.
Vlamingen zijn scherpzinniger dan allochtonen.	Flemish people are sharper than immigrants.
Vlamingen zijn minder dom dan allochtonen.	Flemish people are less dumb than immigrants.
Allochtonen zijn niet zo intelligent als Vlamingen.	Immigrants are not as intelligent as Flemish people.
Allochtonen zijn dommer dan Vlamingen.	Immigrants are dumber than Flemish people.
Allochtonen zijn minder verstandig dan Vlamingen.	Immigrants are less wise than Flemish people.
Allochtonen zijn minder scherpzinnig dan Vlamingen.	Immigrants are less sharp than Flemish people
Allochtonen zijn achterlijker dan Vlamingen.	Immigrants are more backward than Flemish people.
Vlamingen zijn niet zo intelligent als allochtonen.	Flemish people are not as intelligent as immigrants.
Vlamingen zijn dommer dan allochtonen.	Flemish people are dumber than immigrants.
Vlamingen zijn minder verstandig dan allochtonen.	Flemish people are less wise than immigrants.
Vlamingen zijn minder scherpzinnig dan allochtonen.	Flemish people are less sharp than immigrants.
Vlamingen zijn achterlijker dan allochtonen.	Flemish people are more backward than immigrants.
Allochtonen zijn intelligenter dan Vlamingen.	Immigrants are more intelligent than Flemish people.
Allochtonen zijn slimmer dan Vlamingen.	Immigrants are smarter than Flemish people.
Allochtonen zijn verstandiger dan Vlamingen.	Immigrants are wiser than Flemish people.
Allochtonen zijn scherpzinniger dan Vlamingen.	Immigrants are sharper than Flemish people.
Allochtonen zijn minder dom dan Vlamingen.	Immigrants are less dumb than Flemish people.

CHAPTER 5

BODY DISSATISFACTION REVISITED: THE IMPORTANCE OF BELIEFS CONCERNING BOTH ACTUAL AND IDEAL BODY IMAGE [1]

ABSTRACT

The capability of the Relational Responding Task (RRT) as an implicit measure of beliefs was tested in the context of body dissatisfaction. Body dissatisfaction (i.e., negative attitudes towards one's own body) is assumed to originate from a perceived discrepancy between the actual and ideal body image. We assessed implicit beliefs concerning actual and ideal body image using two RRTs. Body dissatisfaction was assessed by means of a (self-report) questionnaire. As hypothesized, participants who exhibited a high level of dissatisfaction with their body exhibited a higher implicit desire to be thin relative to participants who were less dissatisfied with their body. In contrast, the degree to which participants exhibited the implicit belief to be thin was lower in participants who exhibited a high degree of body dissatisfaction as compared to participants who exhibited a low degree of body dissatisfaction. In contrast to previous research, the findings indicated that implicit beliefs concerning both actual and ideal body image contributed to a person's level of self-reported body dissatisfaction. Moreover, the RRT measures allowed for a better identification of body dissatisfied people than the objective index of body mass alone. The findings contribute to the empirical validation of the RRT as an implicit measure of beliefs, as the two applied RRTs differentiated between beliefs concerning actual and ideal body image by focusing on the relational component that distinguishes these beliefs (i.e., a person's *belief* vs. *desire* to be thin or overweight).

[1] Based on Heider, N., Spruyt, A., & De Houwer, J. (2015). Body dissatisfaction revisited: The importance of beliefs concerning both actual and ideal body image. *Manuscript submitted for publication.*

INTRODUCTION

Body dissatisfaction can be defined as the negative attitude towards the own body resulting from a (perceived) discrepancy between one's appearance (i.e., the actual body image) and internalized physical ideals (i.e., the ideal body image; e.g., Cash & Szymanski, 1995; Higgins, 1989; Strauman, Vookles, Berenstein, Chaiken, & Higgins, 1991; Williamson, Gleaves, Watkins, & Schlundt, 1993). Body dissatisfaction has been identified as one of the key factors for dieting behavior, negative affect, and the causation and maintenance of eating disorders (American Psychiatric Association, 2013; Fairburn & Harrison, 2003; Stice, 2001, 2002). Consequentially, since the 1970s, a variety of measures has been developed to assess body dissatisfaction and attitudes related to body image and perceptions (Allebeck, Hallberg, & Espmark, 1976; Bessenoff & Sherman, 2000; Bluemke & Friese, 2012; Degner & Wentura, 2009; Freeman, Thomas, Solyom, & Hunter, 1984; Heider, Spruyt, & De Houwer, 2015; Juarascio et al., 2011; Parling, Cernvall, Stewart, Barnes-Holmes, & Ghaderi, 2012; Roddy, Stewart, & Barnes-Holmes, 2010, 2011; Schlundt & Johnson, 1990; Slade & Russell, 1973).

Until recently, body dissatisfaction was almost exclusively assessed via self-report measures like questionnaires. However, research has shown that explicit measures in general can be biased by social desirability and impression management effects (Cronbach, 1990; Holtgraves, 2004). Moreover, explicit measures have shown to be unsuited to capture attitudes that are introspectively unidentified (Greenwald & Banaji, 1995). Accordingly, researchers developed a new class of diagnostic instruments that allows for an assessment of cognitions in an indirect way, that is, without having to rely on direct self-assessments. Instead, inter-individual differences are inferred from the respondents' response patterns in well-controlled computer tasks, often referred to as *implicit measures* (De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). Expectably, implicit measures have also gained popularity in the field of body related attitudes (e.g., Ahern, Bennett, & Hetherington, 2008; Ahern & Hetherington, 2006; Bluemke & Friese, 2012; Heider et al., 2015; Juarascio et al., 2011). Implicit measures can be divided into two groups. *Associative* implicit measures, like the (standard) evaluative priming task (EPT; Fazio, Jackson, Dunton, & Williams, 1995), the Implicit Association Test (IAT;

Greenwald, McGhee, & Schwartz, 1998), and the affect misattribution paradigm (AMP; Payne, Cheng, Govorun, & Stewart, 2005), are designed to assess the associative strength between concepts in memory (Hughes, Barnes-Holmes, & Vahey, 2012). These measures can, for instance, be used to capture the extent to which a certain class of stimuli (e.g., spiders) is associated with a positive or negative valence. In many cases, however, it is not only important to examine whether two concepts are related in memory but also the precise way in which they are related. In such cases, the application of *relational* implicit measures, like the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2006) and the Relational Responding Task (RRT; De Houwer, Heider, Spruyt, Roets, & Hughes, 2015) seem more appropriate, as they are based on the idea that the processing of propositions is accompanied by an automatic endorsement or rejection (i.e., an automatic evaluation as true or false; e.g., De Houwer, 2014; De Houwer & Moors, 2012; De Houwer et al., 2009). Consequentially, they are capable of assessing *how* relevant concepts are related to another. For the assessment of body dissatisfaction such a sensitivity is beneficial, as body dissatisfaction is characterized as being driven by the perceived discrepancy between one's ideal and actual body image. Crucially, beliefs about these two concepts can be verbalized involving the two concepts self and body size. However, the way in which these two concepts are related to each other differs between the beliefs concerning actual and ideal body, respectively. A belief about one's actual body image involves a relation of descriptiveness to reflect the factuality between the two concepts (e.g., 'I *am* thin'). Beliefs concerning the ideal body image, in contrast, are reflected by a relation of desirability between the self and a certain body size (e.g., 'I *want to be* thin'). The importance of assessing and dissociating between these two types of beliefs was recently demonstrated by Heider et al. (2015, also see Chapter 3). They used one IRAP (hereafter referred to as actual-IRAP) to capture beliefs about the actual body image. A second IRAP (hereafter referred to as ideal-IRAP) was used to capture beliefs about the ideal body image. In the actual-IRAP, participants were presented, on each trial, with combinations of two stimuli that were either congruent or incongruent with possessing a thin body (e.g., 'I am' + 'skinny' and 'I am' + 'chubby', respectively). Crucially, participants were asked to respond as if they believed to be thin in one block of trials, and to respond as if they believed to be overweight in a second block of trials. That is, they were required to select the response 'true' whenever a stimulus combination was presented that reflected the belief participants were asked to respond in accordance to within that block of trials, and to select the response 'false' whenever a stimulus combination was presented that reflected the opposite belief. The ideal-IRAP was identical to the actual-IRAP except for the fact that participants were (a) presented with stimulus combinations that reflected the desire to be thin or overweight (e.g., 'I want to be + skinny' and 'I want to be + chubby', respectively) and (b) required to respond as if they desired to thin or overweight in different blocks of trials.

As mentioned above, the IRAP is based on the idea that the processing of a stimulus combination that expresses a specific belief is accompanied by an automatic endorsement or rejection of that belief. Consequentially, participants are expected to more easily respond to stimuli in a manner that are in line with their personal beliefs than it is to respond in a manner that is inconsistent with their personal beliefs. Therefore, the difference in the performance between the two types of trials within an IRAP can be used to infer a participant's belief. Faster responses in an actual-IRAP on trials that require responding in line with possessing a thin body as compared to on trials that require responding in line with possessing an overweight body for instance indicate the belief of possessing a thin body. Similar inferences about a person's ideal body image (i.e., the desire to possess a thin body) can be made from the differences in response latencies in the two different blocks of trials of the ideal-IRAP.

In accordance with the assumption that body dissatisfaction can be described as a function of a person's actual and ideal body image, Heider et al. (2015) found that participants high and low in body dissatisfaction differed in their scores in the two IRAPs. Specifically, results indicated that the implicit belief to be thin was more pronounced in individuals who were satisfied with their body as compared to individuals that were dissatisfied with their body. The implicit desire to be thin was in contrast more pronounced in individuals who were dissatisfied with their body as compared to individuals who exhibit a low degree of body dissatisfaction. These findings clearly indicate the benefit of assessing beliefs underlying body dissatisfaction as compared to assessing only the strength of association between the concepts of self and body weight (e.g., 'self – thin', 'self – overweight'). Consequentially, we consider these results an important step towards a more precise implicit measurement of body dissatisfaction.

However, there are two reasons that encouraged us to revisit the implicit measurement of beliefs concerning actual and ideal body image in the current article.

The first reason can be found in the additional analyses indicating that the level of body dissatisfaction was dependent upon scores of the ideal-IRAP only. This finding was unexpected, as it was hypothesized that a person's level of body dissatisfaction should also depend on actual body image beliefs. Specifically, based on the definition of body dissatisfaction, Heider et al. (2015) assumed that actual body image beliefs moderated the degree to which the desire to possess a thin body correlates with a person's level of body dissatisfaction. We thus decided to revisit the implicit assessment of actual and ideal body image beliefs with the abovementioned RRT. The RRT is, like the IRAP, based on the idea that propositions are automatically endorsed or rejected when processed. The RRT is therefore sensitive to *how* concepts are related to another. In their seminal study, De Houwer et al. (2015, also see Chapter 4) demonstrated the capabilities of the RRT in assessing racist beliefs. Specifically, participants were asked to respond to propositions that expressed the belief that Flemish people were either more or less intelligent than immigrants (e.g., 'Flemish people are smarter than immigrants', 'Flemish people are less intelligent than immigrants', respectively). In one crucial block of trials, participants were asked to categorize these propositions as true or false in line with believing that Flemish people are more intelligent than immigrants. In a second crucial block of trials, participants were asked to reverse their responses and to categorize the same propositions in accordance with the belief that immigrants are more intelligent than Flemish people.[2] The difference in performance between blocks of different response rules provided a score of the implicit belief that Flemish people are more intelligent than immigrants. Crucially, De Houwer et al. (2015) found that this score correlated with the participants' self-reported endorsement of the belief that Flemish people are more intelligent that immigrants, as well as with multiple more general questionnaire measures of racism.

[2] On additional trials (i.e., inducer trials) throughout the task, participants were asked to categorize synonyms of the response options 'true' and 'not true' in accordance with their meaning. The function of these trials was to discourage participants to recode the responses in term of their physical location.

The second motivation for the present research concerns the RRT itself, namely the claim that it is sensitive to relational information. As propositions concerning actual and ideal body image differ only in how the concepts self and body size are related, this research domain is an ideal test case for the RRT. If the RRT is indeed sensitive to relational information, we should be capable of dissociating between these two beliefs and how they are related to body dissatisfaction.

For these two reasons, in the present study, we addressed beliefs concerning actual and ideal body image using two versions of the RRT (i.e., actual-RRT and ideal-RRT, respectively). The actual-RRT was used to assess beliefs about the actual body and included statements like 'I possess a slim body' and 'I see myself as a fat person'. The ideal-RRT was used to assess beliefs about the desired body ideal and included statements like 'I wish I was thinner' and 'I strive to weigh more'. We hypothesized that participants would differ in the scores of the two RRTs as a function of their level of explicitly measured body dissatisfaction. Specifically, based on our previous findings, we expected the scores of the ideal-RRT to be the main predictor of body dissatisfaction. However, in line with the general definition of body dissatisfaction, we furthermore expected the scores of the actual-RRT to moderate the relationship between ideal body image and body dissatisfaction. Specifically, we expected the thin ideal body image to correspond with high body dissatisfaction only when participants possessed an overweight actual body image.

METHOD

Participants

During an online screening study that involved several questionnaires, 468 students at Ghent University completed the body dissatisfaction subscale of the Eating Disorders Inventory (EDI; Garner et al., 1983) online. About one month later, we invited all female students who had scored within the first and forth quartile of the sample (*N* = 143), using an online participant recruitment system. In total, 68 female students (*M* = 18.72 years, *SD* = 2.12) participated in the experiment in exchange for course credit or payment of €7

(n = 6). Four participants were excluded because either their mean response latencies or error rates exceeded our cutoff criterion of 2.5 standard deviations above the grand means of the respective criterion. Based on the EDI scores obtained during the actual test session, the final sample (N = 64) was divided by cluster analysis into a high body dissatisfaction group (n = 31, M = 19.1, SD = 4.6, min = 12, max = 32) and low body dissatisfied group (n = 33, M = 46.4, SD = 5.2, min = 36, max = 54). Groups differed significantly in terms of mean EDI score, $t(62) = 22.35$, $p < .001$. All participants were Dutch speakers and had normal or corrected-to-normal vision.

Measurement

Relational Responding Task. Ten words were used as stimuli on the inducer trials of the actual-RRT and the ideal-RRT. Five words were related to 'true' and five words were related to 'not true' (for the list of the Dutch words and their English translations, see Table 2 in the Appendix). Two separate sets of 20 statements were used for the target trials of the two RRTs. Statements for the actual-RRT related thin or overweight in a way that concerned the perceived state of the participants' body. Statements in the ideal-RRT concerned the desired state of the participants' body. Crucially, each of the twenty target statements of each RRT contained words or phrases related to thinness or overweight in a self-relevant sentence. In the actual-RRT, five statements each implied the belief of being thin (e.g., 'I possess a slim body') or overweight (e.g., 'I weigh too much'). Negations in these ten statements led to five statements that implied the belief of not being overweight (e.g., 'I do not weigh too much') or not being thin (e.g., 'I do not possess a slim body'), respectively. In the ideal-RRT, five statements each implied that the desire to be thin (e.g., 'I desire to weigh less') or overweight (e.g., 'I desire to be chubbier'). Negations in these ten statements led to five statements that imply the desire of not being overweight (e.g., 'I don't desire to be chubby') or thin (e.g., 'I don't desire to weigh less'), respectively (for the complete list of the statements and their English translations, see Tables 3 and 4 in the Appendix).

Participants were instructed to categorize words and statements presented on screen as 'true' or 'not true' by pressing the right and left ctrl-key of the keyboard, respectively.

Each RRT consisted of five blocks. In Block 1 (40 trials) each of the ten inducer words was presented 4 times. Participants were asked to categorize the words as referring to 'true' and 'not true'. In Block 2 (40 trials), the 20 target statements were presented twice. Participants were asked to respond in line with possessing a thin actual body image (in the actual-RRT) or a thin ideal body image (in the ideal-RRT). Block 3 (80 trials) consisted of a first and a second block of 40 trials each. In each half, the ten inducer words were presented twice and the 20 target statements were presented once. Participants were asked to respond in accordance with the rules practiced in the two preceding blocks. Block 4 (40 trials) was identical to Block 2 except for a reversal of the statement-response assignments. Participants were now asked to respond in line with possessing an overweight actual body image (in the actual-RRT) or desiring to possess an overweight ideal body image (in the ideal-RRT). Block 5 (80 trials) was identical to Block 3, but participants were asked to respond to target statements in accordance with the response rule learned in Block 4. The order of trials within all blocks was semi-randomized, preventing the same word or phrase to be presented on two consecutive trials.

Each trial started with the presentation of an item (i.e., inducer word or target statement) in the middle of the screen, Tahoma, 28 point font size. Inducer stimuli were presented in white color in both RRTs, target statements were presented in orange (actual-RRT) or blue color (ideal-RRT). Different colors were used to increase awareness of the fact that target statements differed between tasks. Items remained on screen until the correct response was registered. Incorrect responses were indicated by the presentation of a red X (Arial, 72 point font size) below the item until participants gave the correct response. The next trial started 750 ms after registration of the correct response. The RRTs were presented on a 17 inch LCD screen (60 Hz, 1440 × 900 pixels) and were written in Affect 4.0 (Spruyt, Clarysse, Vansteenwegen, Baeyens, & Hermans, 2010).

Self-report measures. The 20 statements used on the target trials of each of the RRTs were rated on a 5-point scale, ranging from *completely disagree* to *completely agree*. Ratings of each group of target statements were aggregated to calculate scores of the participants' explicit actual and ideal body image beliefs, with a score of 1 indicating total agreement with being/wanting to be overweight, whereas a score of 5 indicated total agreement with being/wanting to be thin. Body dissatisfaction was assessed by means of the body dissatisfaction subscale of the Eating Disorder Inventory (EDI, 9 items; Garner et al., 1983). Actual and ideal body image were measured using the female version of the Contour Drawing Rating Scale (CDRS; Thompson & Gray, 1995). The CDRS consists of nine schematic (female) figures of varying sizes ranging from *underweight* (1) to *overweight* (9). Participants completed the CDRS twice, once to indicate their actual body image and once to indicate their ideal body image. Finally, we computed the Body Mass Index (BMI) for each participant using their self-reported weight and height.

Procedure. All participants gave informed consent before participation. Participants completed both RRTs in a counterbalanced order. They then indicated their level of agreement with all statements used on the target trials of both RRT tasks. Afterwards participants completed the EDI and CDRSs for actual and ideal body image and reported their weight and height for the calculation of their BMI. After the experiment participants were debriefed. All participants were tested individually. The experiment took approximately 35 minutes.

RESULTS

Data preparation

For the analysis of the RRT data, only the data of the target trials (i.e., trials containing target statements) of the mixed blocks were analyzed.[3] The data of two participants were excluded from the analysis because their mean reaction times in both tasks (2428 ms and 2629 ms, for the actual-RRT; 2862 ms and 2838 ms, for the ideal-RRT) exceeded our cutoff criterion of 2.5 standard deviations above the grand means of the tasks (Actual-RRT: M = 1292 ms, SD = 369 ms; threshold = 2215 ms; Ideal-RRT: M = 1521 ms, SD = 417 ms; threshold = 2563 ms; see Ratcliff, 1993). Also excluded from the analysis were the data of two other participants whose error rates (27.5 % and 35.0 %) in one of the two tasks exceeded the cutoff criterion of 2.5 standard deviations above the grand mean of that

[3] In contrast to the IAT, this second type of trials (i.e., inducer trials) was included only to prevent response recoding. Note, however, that analyses that included data of inducer trials led to similar results.

task (Actual-RRT: *M* = 11.2 %, *SD* = 5.3 %; threshold = 24.4 %; Ideal-RRT: *M* = 11.3 %, *SD* = 6.7 %; threshold = 28 %; Ratcliff, 1993).[4] The mean reaction time on target trials in the actual-RRT was 1284 ms (*SD* = 341 ms), with participants responding incorrectly on 10.9 % (*SD* = 4.8 %) of the target trials. In the ideal-RRT, participants on average needed 1517 ms (*SD* = 394 ms) to respond on target trials and responded incorrectly on 10.8 % (*SD* = 6 %) of the target trials.

For each participant and each version of the RRT, the raw response latencies were transformed into D_{RRT} scores using the D-algorithm (D1) described by Greenwald, Nosek, and Banaji (2003).[5] Positive D_{RRT} scores of the actual-RRT indicated stronger beliefs of possessing a thin body, whereas negative D_{RRT} scores indicated stronger beliefs of possessing an overweight body. Similarly, positive D_{RRT} scores of the ideal-RRT indicated a stronger desire of possessing a thin body, whereas negative D_{RRT} scores indicated a stronger desire of possessing an overweight body. The D_{RRT} score of the actual-RRT ranged from -0.46 to 0.71, with a mean score of *M* = 0.08 (*SD* = 0.30), which differed from zero, t (63) = 2.07, p = .043, d = 0.37. The D_{RRT} scores of the ideal-RRT ranged from -0.63 to 0.63, with a mean score of *M* = -0.05 (*SD* = 0.33), which did not differ from zero, $t(63)$ = -1.33, p = .189, d = -0.23.

Effects at the group level

RRT measures. To investigate whether the RRTs of actual and ideal body image were dependent upon the degree of body dissatisfaction, the overall D_{RRT} scores were submitted to a 2 (*RRT*: actual vs. ideal) x 2 (*body dissatisfaction*: high vs. low) mixed models ANOVA.[6] As expected, we found a significant interaction of body dissatisfaction and RRT, $F(1, 62) = 8.80$, $p = .004$, $\eta\rho^2 = .12$, indicating that D_{RRT} scores differed in both groups. Participants low in body dissatisfaction scored higher on the actual-RRT than participants high in body dissatisfaction, 0.18 vs. -0.02, $t(62) = 2.72$, $p = .008$, d = 0.68.

[4] Analyses that included all participants did not yield to different results.
[5] Analyses using the D4 scoring algorithm led to similar results.
[6] The order in which the two RRT tasks were completed was counterbalanced across participants. Reassuringly, our critical effect (i.e., the interaction between explicit body dissatisfaction and the type of RRT) was not moderated by this counterbalancing factor, $F(1, 60) = 1.65$, $p = .204$, $\eta\rho^2 = .03$, nor were any other effects, all *F*s < 3.18, all *p*s > .08. We therefore excluded the order factor from the analyses.

Conversely, participants high in body dissatisfaction scored higher on the ideal body image RRT than participants low in body dissatisfaction, 0.00 vs. -0.11, $t(62) = 1.40$, $p = .166$, d = 0.35. This pattern of results indicates that the belief to be thin was more pronounced for participants low in body dissatisfaction as compared to participants high in body dissatisfaction, and the desire to possess a thin body was more pronounced for highly body dissatisfied participants as compared to participants low in body dissatisfaction. In addition, we found a significant main effect of RRT, $F(1, 62) = 6.90$, $p = .011$, $\eta\rho^2 = .10$, indicating higher D_{RRT} scores on the actual-RRT as compared to the ideal-RRT, 0.08 vs. -0.05. No other effects were significant, all *F*s < 1, all *p*s > .475.

Explicit ratings of the target statements of the RRT measures. The explicit ratings of the target statements used in both RRTs were used as dependent variables in a 2 (*type of target statement*: actual body image vs. ideal body image) x 2 (*body dissatisfaction*: high vs. low) mixed model ANOVA. Results mirrored those of the D_{RRT} scores. We found a significant interaction of type of target statement and body dissatisfaction, $F(1, 62) = 135.09$, $p < .001$, $\eta\rho^2 = .69$, indicating that participants high and low in body dissatisfaction differed in their ratings of the two types of target statements. Specifically, participants low in body dissatisfaction rated statements more in line with being thin as compared to participants high in body dissatisfaction, 4.09 vs. 1.93, $t(62) = 7.46$, $p < .001$, d = 1.86. Participants high in body dissatisfaction rated statements more in line with wanting to be thin as compared to participants low in body dissatisfaction, 4.55 vs. 3.42 $t(62) = 12.97$, $p < .001$, d = 3.24. Moreover, within groups, mean ratings of both types of statements differed significantly from the scales mean of 2.5, *t*s > 5.49, *p*s < .001. Additionally, we found a significant main effect of the type of statement, $F(1, 62) = 47.83$, $p < .001$, $\eta\rho^2 = .44$, indicating higher overall ratings on statements about the ideal body image, 4.00 vs 2.97. Finally, we found a significant main effect of group, $F(1, 62) = 49.30$, $p < .001$, $\eta\rho^2 = .44$, indicating that highly body dissatisfied subjects rated overall more extreme than subjects of the low body dissatisfied group, 3.24 vs. 3.75.

Table 1. *Correlations between measures.*

		1	2	3	4	5	6	7	8
1	EDI	-	0.54***	-0.36**	0.28*	0.67***	-0.06	-0.85***	0.80***
2	BMI		-	-0.34**	0.00	0.82***	0.48***	-0.76***	0.62***
3	Actual-RRT			-	0.01	-0.28*	-0.10	0.30**	-0.15
4	Ideal-RRT				-	0.03	-0.36**	-0.22^{+}	0.13
5	Actual-CDRS					-	0.45***	-0.80***	0.70***
6	Ideal-CDRS						-	-0.07	-0.04
7	Actual-RRT (statements)							-	-0.85***
8	Ideal-RRT (statements)								-

Note. Actual = Actual body image; Ideal = Ideal body image; RRT = Relational Responding Task; CDRS = Contour Drawing Rating Scale; EDI = Body dissatisfaction measured with the Eating Disorders Inventory; BMI = Body Mass Index
$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Correlational Analyses

Table 1 provides an overview of all pairwise correlations between the D_{RRT} scores of the two RRTs, the CDRS scores of actual and ideal body image, explicit body dissatisfaction as measured with the EDI, the target statement ratings of the two RRTs, and the BMI. As a result from sampling participants with a high or low level of body dissatisfaction, some variables were not normally distributed. Accordingly, Spearman's rank order correlations were computed. Results showed that while the D_{RRT} scores of the actual-RRT correlated significantly with the explicit measures of actual body image (i.e., actual-CDRS scores, actual-RRT statement ratings, BMI scores), they were unrelated to the explicit measures of ideal body image (i.e., ideal-CDRS scores, ideal-RRT statement ratings). The D_{RRT} scores of the ideal-IRAP in contrast correlated significantly with Ideal-CDRS scores, but were unrelated to explicit measures of actual body image (i.e., actual-CDRS scores, actual-RRT statement ratings, BMI scores). Moreover, D_{RRT} scores of actual- and ideal-RRT did not correlate with one another.

Hierarchical Logistic Regression Analyses

To further scrutinize the relationship between the D_{RRT} scores of the two RRTs and self-reported body dissatisfaction, and to ascertain whether the D_{RRT} scores contributed substantially to the predictability of body dissatisfaction over and above an explicit measure of body size, a hierarchical logistic regression was performed. We compared a model in which the BMI scores were the only predictor of body dissatisfaction group membership (i.e., high and low body dissatisfaction as assessed with the EDI) with a model that included the BMI and the D_{RRT} scores of the two RRTs. The BMI alone was predictive of group membership, $\chi^2(1) = 16.27$, $p < .001$, Nagelkerke $R^2 = .31$, $OR = 1.528$ (71.9 % correct classifications). However, adding the D_{RRT} scores of the two RRTs and their interaction into the model resulted in a significantly better model fit, $\chi^2(3) = 10.35$, $p = .016$. On the basis of the model including the BMI and the D_{RRT} scores of the two RRTs and their interaction, $\chi^2(4) = 27.17$, $p < .001$, Nagelkerke $R^2 = .46$, 76.6 % of all participants were correctly classified. Specifically, both the BMI, $\chi^2(1) = 9.41$, $p = .002$, $OR = 1.580$, and the D_{RRT} scores of the ideal-RRT, $\chi^2(1) = 5.20$, $p = .023$, $OR = 23.160$, contributed significantly to the prediction of group membership. The D_{RRT} scores of the actual-RRT and the interaction term just missed conventional significance levels, $\chi^2(1) = 3.52$, $p = .061$, $OR = .093$, and $\chi^2(1) = 3.71$, $p = .054$, $OR = 0.001$, respectively.[7] Specifically, the interaction indicates that the degree to which a thin ideal body image is predictable of belonging to the group of body dissatisfied participants is dependent on the participant's beliefs about the actual body image. For participants that perceived themselves as overweight, body dissatisfaction is dependent on whether or not a participant desired to be thin. For participants that in contrast saw themselves as thin, the desire to be thin did not influence group membership (see Figure 1).

Reliability

For the EDI subscale of body dissatisfaction, Chronbach's alpha was 0.98. Test-retest reliability of the actual body image CDRS was r = .78, as reported by Thompson & Gray (1995). Reliability coefficients of both RRTs were estimated using a bootstrap procedure, wherein 100 random-splits were drawn from the data. For each random split, the correlation across participants between the two RRT scores was calculated. Correlations were then averaged. This procedure resulted in spearman-brown corrected mean split-half correlations of Rsb = 0.49 and Rsb = 0.57, for the actual- and ideal-RRT, respectively.

[7] Noteworthy, this better model fit depended on the inclusion of the D_{RRT} scores of the actual-RRT. A model containing only the BMI, the D_{RRT} scores of the ideal-RRT and their interaction did not fit the data significantly better than a model containing the BMI alone, $\chi^2(2) = 3.96$, $p = .138$.

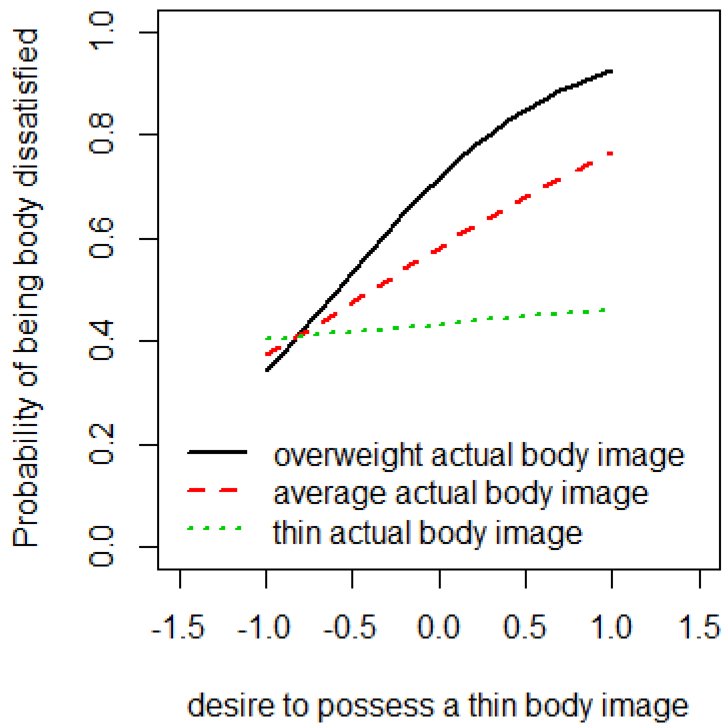

Figure 1. *Probability of membership to group of high body dissatisfied participants as a function of ideal body image D_{RRT} scores, separated for thin (i.e., 1 SD below average), average, and overweight (i.e., 1 SD above average) actual body image D_{RRT} scores.*

DISCUSSION

The degree to which people are dissatisfied with their own body is a function of the (perceived) discrepancy between one's actual and ideal body image (Cash & Szymanski, 1995; Higgins, 1989). We recently demonstrated that young female adults that are either high or low in body dissatisfaction differed not only in their self-reported degree of body dissatisfaction but also in their implicit beliefs concerning their actual and ideal body image, as measured with two IRAPs (Heider et al., 2015). In the present study, we revisited the implicit measurement of body dissatisfaction with the RRT (De Houwer et al., 2015) and replicated our earlier findings.

We again showed that the implicit belief to be thin was more pronounced for participants low in body dissatisfaction as compared to participants high in body dissatisfaction. The implicit desire to be thin was, in contrast, more pronounced in participants high in body dissatisfaction as compared to participants low in body dissatisfaction. We observed this pattern not only for the scores of the RRTs, but also in self-report ratings based on the statements used in the two RRTs. These findings first and foremost suggest that the RRT is capable of dissociating between closely related beliefs that differ only in their relational component (i.e., 'I *am* thin' vs. 'I *want to be* thin'). Additional findings corroborated this inference. While the two RRTs did not correlate with each other, D_{RRT} scores of the actual-RRT correlated with explicit measures of actual body image but not with explicit measures of ideal body image. D_{RRT} scores of the ideal-RRT on the other hand correlated with explicit measures of ideal body image but not with explicit measures of actual body image. Such a finding strongly supports the conclusion that both RRT measures, despite their structural similarity, captured different beliefs.

In addition, the present results go beyond our earlier findings in several ways. First, whereas Heider et al. (2015) observed that the level of body dissatisfaction was dependent upon scores of the ideal-IRAP only, we found that the scores of both implicit measures contributed significantly to the prediction of body dissatisfaction. Second, results hinted that explicit body dissatisfaction was dependent upon the interaction between the two RRT scores. In line with the idea that body dissatisfaction is a function of the (perceived) discrepancy between one's actual and ideal body image, we observed that the relationship between body dissatisfaction and the implicit desire to be thin was moderated by implicit beliefs about the own actual body image.

Taken together, these findings strongly suggest that the RRT might allow for a more precise measurement of implicit beliefs concerning actual and ideal body image than the IRAP. The following factor might account for this observation. Note, that the stimuli used in an IRAP are combinations of sample stimuli (e.g., 'I want to be') and a range of target stimuli that are all quite similar to each other (e.g., 'thin', 'lean', etc.). The RRT in contrast allows for the use of a much wider range of experimental stimuli in the form of propositions (e.g., 'I strive to weigh less', 'It is my wish to be slimmer'). This difference in the restrictiveness of the experimental items might have influenced the processing of the stimuli on two levels. First, it is imaginable that the use of propositions that differ in the structure of the sentences (RRT), as compared to the rigid combination of sample and target items (IRAP), might have required participants in the current study to process the statements more closely on every trial. The stimuli combinations used in the IRAP might have allowed for a more superficial processing. The possibility of a rather shallow processing of the IRAP stimuli could have been furthermore supported given the higher number of overall trials in the IRAP. As a consequence, the likelihood of an automatic endorsement or rejection of the stimuli based on the participant's subjective beliefs (i.e., the mechanism from which RRT effects are expected to originate from; see De Houwer et al., 2015) might have been higher for the RRT as compared to the IRAP. It is also imaginable that the diversity in the statements might have allowed for capturing more fine-grained beliefs with the RRT. Similar to a well-constructed questionnaire, the higher number of different propositions that differ slightly in their wording might have increased the precision of the measure in assessing the specific construct. While both explanations seem intuitively plausible, research is needed to back up these claims and to properly investigate the differences in the two relational implicit measures.

As indicated above, the number of published studies in which the RRT was used as an implicit measure of beliefs is still meager. In their initial demonstration of the capabilities of the RRT as an implicit measure of beliefs, De Houwer et al. (2015) focused on a single belief and examined whether an RRT measure of that belief correlates with criterion variables. In the current study we continued the empirical validation of the RRT by testing its potential in capturing differences between beliefs that vary only in their relational component. As mentioned above, beliefs concerning the actual and ideal body image vary only in how the inherent concepts self and body size are related (i.e., 'am' vs 'want to be', and variations thereof). The finding that both RRT measures were related to different explicit variables to a different extent underscores the distinctiveness of both measures and their sensitivity to relational information. The present findings can thus be seen as a crucial step in the empirical validation process of the RRT.

Furthermore, the current findings support the propositional models of implicit cognition (e.g., De Houwer, 2009, 2014; Mitchell, De Houwer, & Lovibond, 2009) that constitute the theoretical basis of the RRT. Most models of implicit cognition postulate that the automatic processing of information is based solely on the activation of associations in long-term memory (e.g., Collins & Loftus, 1975). By definition, such models do not allow for the implicit endorsement of beliefs, as associative representations do not possess a truth value. In contrast to this predominant view, De Houwer et al. (2015) argued that automatic endorsement of processed propositional knowledge is possible, if one does not exclude the involvement of propositions from the realm of implicit cognition (De Houwer, 2014; Hughes, Barnes-Holmes, & De Houwer, 2011). The presented findings support this assumption, as they indicate that the endorsement or rejection of propositions, including their inherent relational component, can take place in an automatic fashion. At the theoretical level, our results thus give additional credibility to propositional models of implicit cognition.

To conclude, with the present paper we continued the path of empirically validating the RRT as an implicit measure of beliefs, as well as investigating the relationship between body dissatisfaction and beliefs concerning actual and ideal body image. Two RRTs were used to capture and dissociate between the two closely related concepts actual and ideal body image. Results indicated that beliefs concerning both actual and idea body image were related to the level to which a person is satisfied with the own body. More generally, the findings substantiate the claim that the RRT can be used as a tool to capture beliefs at the implicit level.

REFERENCES

Ahern, A. L., Bennett, K. M., & Hetherington, M. M. (2008). Internalization of the Ultra-Thin Ideal: Positive Implicit Associations with Underweight Fashion Models are Associated with Drive for Thinness in Young Women. *Eating Disorders*, *16*, 294–307. doi: 10.1080/10640260802115852

Ahern, A. L., & Hetherington, M. M. (2006). The thin ideal and body image: An experimental study of implicit attitudes. *Psychology of Addictive Behaviors*, *20*, 338–342. doi: 10.1037/0893-164X.20.3.338

Allebeck, P., Hallberg, D., & Espmark, S. (1976). Body image—an apparatus for measuring disturbances in estimation of size and shape. *Journal of Psychosomatic Research*, *20*, 583–584. doi: 10.1016/0022-3999(76)90060-X

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition). Arlington, VA: American Psychiatric Pub.

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, *32*, 169–177.

Bessenoff, G. R., & Sherman, J. W. (2000). Automatic and Controlled Components of Prejudice Toward Fat People: Evaluation Versus Stereotype Activation. *Social Cognition*, *18*, 329–353. doi: 10.1521/soco.2000.18.4.329

Bluemke, M., & Friese, M. (2012). On the Validity of Idiographic and Generic Self-Concept Implicit Association Tests: A Core-Concept Model. *European Journal of Personality*, *26*, 515–528. doi: 10.1002/per.850

Cash, T. F., & Szymanski, M. L. (1995). The Development and Validation of the Body-Image Ideals Questionnaire. *Journal of Personality Assessment*, *64*, 466–477. doi: 10.1207/s15327752jpa6403_6

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428. doi: 10.1037/0033-295X.82.6.407

Cronbach, L. (1990). *Essentials of psychological testing*. New York, NY: Harper & Row.

Degner, J., & Wentura, D. (2009). Not Everybody Likes the Thin and Despises the Fat: One's Weight Matters in the Automatic Activation of Weight-Related Social Evaluations. *Social Cognition*, *27*, 202–221. doi: 10.1521/soco.2009.27.2.202

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*, 1–20. doi: 10.3758/LB.37.1.1

De Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass*, *8*, 342–353. doi: 10.1111/spc3.12111

De Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The Relational Responding Task: Toward a New Implicit Measure of Beliefs. *Frontiers in Psychology*, *6:319*. doi: 10.3389/fpsyg.2015.00319

De Houwer, J., & Moors, A. (2012). How to define and examine implicit processes. In R. Proctor & J. Capaldi (Eds.), *Psychology of science: Implicit and explicit processes* (pp. 183–198). Oxford: Oxford University Press.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368. doi: 10.1037/a0014211

Fairburn, C. G., & Harrison, P. J. (2003). Eating disorders. *The Lancet*, *361*, 407–416. doi: 10.1016/S0140-6736(03)12378-1

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027. doi: 10.1037/0022-3514.69.6.1013

Freeman, R. J., Thomas, C. D., Solyom, L., & Hunter, M. A. (1984). A modified video camera for measuring body image distortion: Technical description and reliability. *Psychological Medicine*, *14*, 411–416. doi: 10.1017/S0033291700003652

Garner, D. M., Olmstead, M. P., & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders*, *2*, 15–34. doi: 10.1002/1098-108X(198321)2:2<15::AID-EAT2260020203>3.0.CO;2-6

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27. doi: 10.1037/0033-295X.102.1.4

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216. doi: 10.1037/0022-3514.85.2.197

Heider, N., Spruyt, A., & De Houwer, J. (2015). Implicit Beliefs Concerning Ideal Body Image Predict Body Dissatisfaction. *Manuscript Submitted for Publication*.

Higgins, E. T. (1989). Self-Discrepancy Theory: What Patterns of Self-Beliefs Cause People to Suffer? *Advances in Experimental Social Psychology*, *22*, 93–136. doi: 10.1016/s0065-2601(08)60306-8

Holtgraves, T. (2004). Social Desirability and Self-Reports: Testing Models of Socially Desirable Responding. *Personality and Social Psychology Bulletin*, *30*, 161–172. doi: 10.1177/0146167203259930

Hughes, S., Barnes-Holmes, D., & De Houwer, J. (2011). The dominance of associative theorizing in implicit attitude research: Propositional and behavioral alternatives. *The Psychological Record*, *61*, 464–498.

Hughes, S., Barnes-Holmes, D., & Vahey, N. (2012). Holding on to our functional roots when exploring new intellectual islands: A voyage through implicit cognition research. *Journal of Contextual Behavioral Science*, *1*, 17–38. doi: 10.1016/j.jcbs.2012.09.003

Juarascio, A. S., Forman, E. M. ., Timko, C. A., Herbert, J. D., Butryn, M., & Lowe, M. (2011). Implicit internalization of the thin ideal as a predictor of increases in weight, body dissatisfaction, and disordered eating. *Eating Behaviors*, *12*, 207–213. doi: 10.1016/j.eatbeh.2011.04.004

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*, 183–198. doi: 10.1017/S0140525X09000855

Parling, T., Cernvall, M., Stewart, I., Barnes-Holmes, D., & Ghaderi, A. (2012). Using the Implicit Relational Assessment Procedure to Compare Implicit Pro-Thin/Anti-Fat Attitudes of Patients With Anorexia Nervosa and Non-Clinical Controls. *Eating Disorders*, *20*, 127–143. doi: 10.1080/10640266.2012.654056

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*, 277–293. doi: 10.1037/0022-3514.89.3.277

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*, 510–532. doi: 10.1037/0033-2909.114.3.510

Roddy, S., Stewart, I., & Barnes-Holmes, D. (2010). Anti-fat, pro-slim, or both?: Using two reaction-time based measures to assess implicit attitudes to the slim and overweight. *Journal of Health Psychology*, *15*, 416–425. doi: 10.1177/1359105309350232

Roddy, S., Stewart, I., & Barnes-Holmes, D. (2011). Facial reactions reveal that slim is good but fat is not bad: Implicit and explicit measures of body-size bias: Facial electromyography and body-size bias. *European Journal of Social Psychology*, *41*, 688–694. doi: 10.1002/ejsp.839

Schlundt, D. G., & Johnson, W. G. (1990). *Eating disorders: Assessment and treatment.* Needham Heights, MA: Allyn & Bacon.

Slade, P. D., & Russell, G. (1973). Awareness of body dimensions in anorexia nervosa: Cross-sectional and longitudinal studies. *Psychological Medicine*, *3*, 188–199. doi: 10.1017/S0033291700048510

Spruyt, A., Clarysse, J., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2010). Affect 4.0: A Free Software Package for Implementing Psychological and Psychophysiological Experiments. *Experimental Psychology (formerly Zeitschrift Für Experimentelle Psychologie)*, *57*, 36–45. doi: 10.1027/1618-3169/a000005

Stice, E. (2001). A prospective test of the dual-pathway model of bulimic pathology: Mediating effects of dieting and negative affect. *Journal of Abnormal Psychology*, *110*, 124–135. doi: 10.1037/0021-843X.110.1.124

Stice, E. (2002). Risk and maintenance factors for eating pathology: a meta-analytic review. *Psychological Bulletin*, *128*, 825–848. doi: 10.1037/0033-2909.128.5.825

Strauman, T. J., Vookles, J., Berenstein, V., Chaiken, S., & Higgins, E. T. (1991). Self-discrepancies and vulnerability to body dissatisfaction and disordered eating. *Journal of Personality and Social Psychology*, *61*, 946–956. doi: 10.1037/0022-3514.61.6.946

Thompson, M. A., & Gray, J. J. (1995). Development and Validation of a New Body-Image Assessment Scale. *Journal of Personality Assessment*, *64*, 258–269. doi: 10.1207/s15327752jpa6402_6

Williamson, D. A., Gleaves, D. H., Watkins, P. C., & Schlundt, D. G. (1993). Validation of self-ideal body size discrepancy as a measure of body dissatisfaction. *Journal of Psychopathology and Behavioral Assessment*, *15*, 57–68. doi: 10.1007/BF00964324

APPENDIX

Table 2. *Inducer words of the RRTs and English translations*

Inducer word	English translations	Inducer word	English translations
goed	good	Mis	wrong
juist	right	Onjuist	incorrect
correct	correct	incorrect	incorrect
exact	exact	verkeerd	wrong
in orde	alright	Fout	error

Table 3. *Target statements of the actual-RRT and English translations*

Target statements	English translations
Ik heb een tengere lichaamsbouw.	I have a slender physique.
Ik weeg weinig.	I weigh little.
Ik zie mezelf als een slank persoon.	I consider myself a slim person.
Ik heb een mager figuur.	I have a lean figure.
Ik ben fijngebouwd.	I am delicately built.
Ik heb geen zware lichaamsbouw.	I do not have a plump physique.
Ik weeg niet te veel.	I do not weigh too much.
Ik zie mezelf niet als een dik persoon.	I do not consider myself a fat person.
Ik heb geen mollig figuur.	I do not have a stout figure.
Ik ben niet struis.	I am not robustly built.
Ik heb geen tengere lichaamsbouw.	I do not have a slender physique.
Ik weeg niet te weinig.	I do not weigh too little.
Ik zie mezelf niet als een slank persoon.	I do not consider myself a slim person.
Ik heb geen mager figuur.	I do not have a lean figure.
Ik ben niet fijngebouwd.	I am not delicately built.
Ik heb een zware lichaamsbouw.	I have a plump physique.
Ik weeg te veel.	I weigh too much.
Ik zie mezelf als een dik persoon.	I consider myself a fat person.
Ik heb een mollig figuur.	I have a stout figure.
Ik ben struis.	I am robustly built.

Table 4. *Target statements of the ideal-RRT and English translations*

Target statements	English translations
Ik wil een meer tengere lichaamsbouw hebben.	I want to have a more slender physique.
Ik streef ernaar minder te wegen.	I strive to weigh less.
Het is mijn wens om slanker te zijn.	It is my wish to be slimmer.
Ik wil een mager figuur hebben.	I want to have a leaner figure.
Ik wil fijner gebouwd zijn.	I want to be more delicately built.
Ik wil een minder zware lichaamsbouw hebben.	I want to have a less plump physique.
Ik streef er niet naar meer te wegen.	I don't strive to weigh more.
Het is niet mijn wens om dikker te zijn.	It is not my wish to be fatter.
Ik wil geen mollig figuur hebben.	I don't want to have stouter figure.
Ik wil minder struis zijn.	I want to be less robustly built.
Ik wil een minder tengere lichaamsbouw hebben.	I want to have a less slender physique.
Ik streef er niet naar minder te wegen.	I don't strive to weigh less.
Het is mijn wens om minder slank te zijn.	It is my wish to be less slim.
Ik wil geen mager figuur hebben.	I don't want to have a leaner figure.
Ik wil minder fijn gebouwd zijn.	I want to be less delicately built.
Ik wil een zwaardere lichaamsbouw hebben.	I want to have a plumper physique.
Ik streef ernaar meer te wegen.	I strive to weigh more.
Het is mijn wens om dikker te zijn.	It is my wish to be fatter.
Ik wil een mollig figuur hebben.	I want to have stouter figure.
Ik wil struiser zijn.	I want to be more robustly built.

CHAPTER 6

GENERAL DISCUSSION

RECAPITULATION OF THE RESEARCH GOALS

The research presented in this dissertation was inspired by the idea that propositions, rather than associations, guide automatic behavior, as proposed by single-process propositional accounts of cognition (De Houwer, 2009, 2014; Mitchell, De Houwer, & Lovibond, 2009). This view contrasts sharply with predominant dual-process models of cognition as these models postulate that automatic behavior must be guided by processes that are associative in nature. The focus of my dissertational research was twofold. First, it was investigated whether propositions can be constructed under conditions of automaticity. While such an observation would be fully compatible with a single-process propositional model, it would very be hard, if not impossible, to reconcile with classic dual-process models. Second, the usefulness and scope of implicit measures was reexamined in light of the observation that propositional processes can occur under conditions of automaticity. Specifically, it was tested whether the endorsement of propositions can occur under conditions of automaticity. In addition, it was examined whether this logic can be exploited as a means to indirectly assess beliefs (i.e., propositions that are believed to be true). The results presented in each chapter are outlined in detail in the following section.

SUMMERY AND DISCUSSION OF THE RESULTS

Research line 1: The demonstration of automatic propositional processing

In **Chapter 1**, I set out to investigate the abovementioned idea that propositional representations, i.e., informational units in memory that contain information about how concepts are related (De Houwer, 2009, 2014), can be constructed or derived under conditions of automaticity. Specifically, it was tested whether participants are capable of automatically processing the size relation between two objects (i.e., automatic relational stimulus processing). To answer this question, the sequential priming paradigm was used. The behavioral outcome of priming tasks (i.e., the priming effect) allows for inferences about the knowledge that is automatically constructed when processing the irrelevant prime stimuli. Accordingly, in my studies, the priming task was used to examine whether the processing of prime stimuli can lead to an automatic construction of propositional knowledge. Four sequential priming experiments were conducted. In each experiment, participants were asked to judge the physical size of a series of target objects relative to either a large reference object (i.e., a car) or a small reference object (i.e., a football). Crucially, the set of prime objects included objects that were smaller than the large reference object and larger than the small reference object (e.g., a bike). Results showed that the impact on target categorization of these medium-sized objects was contingent upon the size of the reference object. Specifically, it was observed that the medium-sized primes facilitated responses towards large targets and inhibited responses towards small targets when the size of the reference object was small. Conversely, when the size of the reference object was large, the same set of medium-sized primes facilitated responses towards small targets and inhibited responses towards large targets. This interaction, termed relational priming effect, was obtained when the size of the reference object was manipulated block-wise (Experiments 1 & 3), trial-wise (Experiments 2 & 4), and even when the primes were presented near subjective recognition thresholds (Experiment 4). The consistent finding of the relational priming effect implies that participants processed the size of the primes in relation to the size of the reference object of the trial.

To understand why such a finding is problematic for processing models that explain automatic stimulus processing solely in terms of associative processes (i.e., dual-process models), two questions have to be discussed separately. First, why is relational processing of primes considered to be a propositional process? Second, in what way can relational processing of primes be considered to occur automatically?

With regard to the first question, consider how the processing of medium-sized prime stimuli is thought of in a dual-process model. When a specific prime, for instance the prime 'bike', is perceived, its mental representation is activated within an associative network. The activation then spreads automatically via associative links to related concepts (e.g., Fazio, 2001; Strack & Deutsch, 2004). If relational information is also represented within the associative system, one can hypothesize that the two relations of size 'larger than' and 'smaller than' are also associatively connected to the concept 'bike'. However, the relational priming effect implies an activation of only one and not the other relation, depending on the reference object. Such a processing cannot be explained within a purely associative network. In dual-process models, such (relational) processing is therefore assumed to rely on propositional processing, termed propositional categorization (Strack & Deutsch, 2004). Crucially, propositional categorizations require the use of propositional beliefs. For instance, the construction of the proposition "a bike is smaller than a car", requires the retrieval of the two relevant concepts (e.g., the prime bike, the reference object car) and the relation (e.g., smaller than) that is consistent with the truth value of the to-be-constructed proposition from the associative system. The process required for the assembly of the proposition are therefore not purely associative. Moreover, this assembly process is assumed to be time-consuming and to depend upon cognitive resources (e.g., Strack & Deutsch, 2004), and is therefore excluded from occurring under automaticity conditions. This leads to the second question. In what way does the processing of the prime stimuli in the size-judgment task operate automatically? To discuss this question, it is important to understand that a process does not simply operate automatically or not. Automaticity can be decomposed into multiple independent automaticity features, like speed, efficiency, uncontrollability, and unconsciousness (Moors, 2016; see Moors & De Houwer, 2006 for a more extensive list). Whether or not processing occurs given a specific automaticity feature is determined by a multitude of factors like personal characteristics (e.g., personality, learning history, genetic makeup) and the specific experimental procedures used to examine it (e.g., the specific stimulus set, the measure that is used; for an in-depth discussion and a systematic taxonomy of possible factors, see Moors, 2016). Given this decompositional view of automaticity, the automaticity of the relational processing has to be considered with regard to all automaticity features separately.

First, a process is called fast if it consumes little time or has a short duration (Moors, 2016). In the size-judgment priming tasks, both the time between prime and target presentation and the presentation time of the primes were quite short in each of the four experiments. Further, the speed with which responses were made was comparable to that in other implicit measures. Relational prime processing can therefore be considered a fast-acting process. Second, a process is called efficient if it can operate without (or little) cognitive capacities (Moors, 2016). The response task implemented in the size-judgment priming tasks can be considered cognitively more demanding than the response tasks typically used in sequential priming research (e.g., identification of target valence). It can thus be argued that the RPE emerged under conditions of reduced cognitive capacity, implying that the relational processing of the primes driving the RPE operates efficiently. Third, if a process is not affected by either the goal to initiate that process or the goal to counteract it, the process occurs in an uncontrolled fashion (Moors, 2016). It seems unlikely that participants possessed the (conscious) goal to intentionally process the size relation between primes and reference object. Participants were asked to respond on the basis of the targets, and processing primes in the same way was neither informative nor beneficial for the task at hand. However, it is obvious that participants had the goal to process target stimuli with regard to their size relative to the reference objects. As this specific processing goal was crucial for the relational processing of the prime stimuli, the processing of primes can be considered unintentional but goal-dependent. Last, unconsciousness of a process denotes the fact to be unaware of the stimulus input, the process itself, or how the process affects behavior (Moors, 2016). The question whether conscious awareness of the stimulus input is necessary for the RPE to occur was answered inconclusively in Experiment 4. Although prime stimuli were masked and presented only briefly, they were on average still perceived by most participants, as indicated by prime visibility analyses. Accordingly, the cumulative evidence allows to conclude that relational processing can take place under at least three automaticity conditions (see Moors, 2016; Moors & De Houwer, 2006).

On the basis of this discussion of the nature of relational processing and the automaticity conditions under which it can operate, it can be concluded that relational stimulus processing is an automatic propositional process. The observation of relational priming effects in four experiments implies that propositional knowledge can be constructed under conditions of automaticity. The first research goal was achieved.

Research line 2: The indirect assessment of beliefs

The second goal of my dissertational research was to develop and empirically validate an implicit measure of beliefs. As indicated by the findings of Chapter 2, automatic behavior can be guided by propositional processes. This finding led to the idea that the processing of propositional statements might be accompanied by automatic endorsement or rejection (i.e., the evaluation of the truth value) of the statement. If so, an indirect measurement procedure is imaginable that capitalizes on this process to indirectly assess beliefs (i.e., propositions that are believed to be true). The second part of this dissertation (i.e., Chapters 3 to 5) was devoted to the development and empirical validation of such an implicit measure.

The psychological construct of body dissatisfaction played an important role in this endeavor, as its measurement was expected to profit from the indirect measurement of beliefs. Body dissatisfaction can be seen as the negative attitude towards one's own body resulting from a (perceived) discrepancy between one's actual physical appearance (i.e., the actual body image) and internalized ideals about one's physical appearance (i.e., the ideal body image; e.g., Cash & Szymanski, 1995; Higgins, 1989; Strauman, Vookles, Berenstein, Chaiken, & Higgins, 1991; Williamson, Gleaves, Watkins, & Schlundt, 1993). Both constructs, actual and ideal body image, are very dissimilar with respect to how the concepts self and body size are related. Beliefs about the actual body image are defined by a descriptive relation between the self and body size (e.g., "I *am* thin"), whereas beliefs about the ideal body image are defined in terms of desirability (e.g., “I *want to be* thinner”).

In **Chapter 3**, both beliefs were assessed independently with two versions of the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2006) to test whether body dissatisfaction is dependent on implicit beliefs concerning the actual and ideal body image. It was hypothesized that participants high and low in body dissatisfaction would differ both in their self-reported degree of body dissatisfaction and their implicit beliefs concerning these two types of body image. The implicit belief to be thin was expected to be stronger in participants low in body dissatisfaction as compared to participants high in body dissatisfaction. The implicit desire to be thin was in contrast expected to be stronger in participants who exhibit a high degree of body dissatisfaction as compared to participants low in body dissatisfaction. Both predictions were supported by the data of the two IRAP measures. Moreover, the pattern of correlations between each of the two IRAP measures and a number of other explicitly assessed variables was different. Based on these findings, it was concluded that both IRAP measures reflected different underlying constructs, i.e., beliefs about actual and ideal body image, respectively.

Next, I engaged in the development and validation of the Relational Responding Task (RRT). The RRT is a reaction time task that was designed to capitalize on the fact that propositional statements, when processed, are automatically endorsed or rejected (i.e., evaluated as true or false) on the basis of their (subjective) truth value. If a person, for instance, is presented with the propositional statement “a banana is larger than a car”, that person is expected to automatically reject this proposition (i.e., evaluate it as false) based on the knowledge that the opposite is true (i.e., that a banana is smaller than a car). Respondents in an RRT are presented with propositional statements and are asked to respond *as if* they believe a certain type of proposition to be true (e.g., that pieces of fruit are smaller than vehicles) or false (e.g., that vehicles are smaller than pieces of fruit). Performance in these two task-conditions should reflect beliefs about the size of pieces of fruit in relation to vehicles. Specifically, it should be easier to categorize propositions as true and false in a manner that is in line with one's personal beliefs than it is to respond in a manner that is inconsistent with one's personal beliefs.[1]

In **Chapter 4**, such a finding (i.e., better performance in the categorization of statements as true and false when performed in line with one's own beliefs) was obtained in the context of racist beliefs. Specifically, participants were asked to categorize statements about the intelligence of Flemish people relative to immigrants (e.g., "Flemish people are smarter than immigrants") as true or false in accordance with the belief that Flemish people are more intelligent than immigrants in one block of trials, but in accordance with the contrary belief (i.e., that immigrants are more intelligent than Flemish people) in a second block of trials. The difference in the performance between the two types of trials was indeed found to be informative about the participant's racist belief, as the difference in performance between blocks correlated with ratings of the extent to which participants explicitly endorsed the belief that Flemish people are more intelligent than immigrants and with questionnaire measures of subtle and blatant racism.

To further strengthen the argument that propositions can be automatically endorsed or rejected when processed, in **Chapter 5,** the assessment of body dissatisfaction was revisited. Beliefs about actual and ideal body image of young female adults were assessed separately with two RRTs. Reassuringly, RRT effects indicated that both beliefs did not only differ from each other, but also contributed significantly and uniquely to the prediction of the explicitly assessed level of body dissatisfaction of participants. The dissociation of two beliefs that differed only with respect to their relational information was a crucial step in the validation process of the RRT, as it demonstrated that the RRT is capable of capturing differences between beliefs that vary only in their relational component.

To summarize, in the second part of my dissertation, experimental evidence was presented that the subjective truth value of propositional statements can be activated under conditions of automaticity. This conclusion is based on the findings obtained with

[1] The example with statements about the size of pieces of fruit and vehicles was taken from an unpublished pilot study of the RRT. A short summary of the results can be found in Appendix A (Experiment 7).

two implicit measures, the IRAP and the RRT. Both measurement procedures capitalize on the fact that people are generally faster in categorizing statements as true and false in accordance with their own beliefs as compared to in contrast to their own beliefs. In each of these studies the score of the implicit measure correlated with explicit criterion variables of the to-be-measured belief.

To define the endorsement of propositional knowledge as an automatic propositional process, two points need to be discussed, namely why the endorsement of propositions can be considered to be propositional in nature, and under what conditions of automaticity it can operate. According to dual-process models, the endorsement of propositional knowledge is based on propositional processing. Whether a proposition is endorsed or not depends on the consistency of the proposition with other propositional knowledge (i.e., it is based on the principle of preserving consistency between elements within the propositional system; e.g., Gawronski & Strack, 2004; Strack & Deutsch, 2004). For instance, the statement “I want to be thinner” is endorsed if it corresponds with a person’s beliefs about his or her ideal body image. It is rejected if it is in contrast to that belief. Crucially, reasoning processes that involve the comparison of propositional knowledge are propositional in nature. In dual-process models, such processes are assumed to occur only slowly and to require cognitive resources (e.g., Strack & Deutsch, 2004). Single-process propositional accounts, in contrast, propose that propositional processing can also occur automatically. To argue that the endorsement of propositions can occur automatically, the automaticity of the process has to be considered with regard to the separate automaticity features (e.g., speed, efficiency, uncontrollability, consciousness; Moors, 2016). Note that as my research focused mainly on the RRT, the discussion of the automaticity of the endorsement of propositions will focus on the automaticity conditions realized in this task. Further, future research will be proposed regarding those conditions of automaticity for which the research presented in Chapters 4 and 5 is not sufficient for a conclusive answer.

First, consider the automaticity feature of speed. Reaction times to statements in the RRT were undoubtedly longer than reaction times in other implicit measures. However, it is not claimed that the statements themselves were processed under conditions of automatically. The process of interest is the processing of the relation of the statements to their truth values. The relevant question therefore is whether this process operates fast. Considering that participants had to process the complete statements before being able to judging them as true or false in accordance with the response rule, responses can be considered to be relatively fast. For this conclusion, also consider the inducer trials in the RRT. On these trials participants categorized single items as synonyms of true and false, a task that is more comparable to that of other implicit measures. As the speed on those trials was comparable to that in other implicit measures, it can be argued that participants always responded as fast as possible. With regard to the target statements, this means participants responded as soon as the stimuli were completely processed and understood. Consequentially, it can be assumed that the endorsement of statements in line with the personal belief occurred fast, once stimuli themselves were processed with regard to their meaning. However, to substantiate the claim that the endorsement of target statements in an RRT occurs in a fast manner, it might be reasonable to experimentally assess the speed with which statements were processed (in terms of being understood). To do so, one could conduct an experiment in which response latencies to statements in an RRT are compared to response latencies to the same statements in a task in which participants do not have to endorse the statements. Participants in such a task could in contrast be asked to simply indicate the point in time at which they finished processing the meaning of the statements. Comparing response times between these two tasks would allow for an inference about the average time it takes participants to endorse or reject a statement in line with the response rule, once they have processed it. As the endorsement of the statements in line with the personal beliefs (i.e., the process in question) can only start once a statement is understood, such an experiment would allow narrowing down the time window within which the personal endorsement of statements operates.

Second, if a process is not affected by a manipulation of cognitive load, it is considered automatic in terms of efficiency (e.g., De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009). A first step in investigating this issue for the RRT could be to manipulate the working memory load of participants that perform an RRT. Participants could, for instance, be asked to recite a series of digits (Hermans, Crombez, & Eelen, 2000) on a subset of RRT trials. The finding that such a manipulation of the cognitive load does not decrease the magnitude of the RRT effects as compared to the control trials (i.e., trials without the additional digit-reciting task) would indicate that the endorsement of statements is automatic in terms of being efficient. Note, that it can be expected that overall reaction times will increase if participants are asked to perform a second task. This is, because the processing and understanding of the target statements in itself is more demanding than the processing of single stimuli. However, as argued above, of crucial interest is not the processing of the target statements, but their endorsement in line with personal beliefs. Therefore, of crucial interest is whether or not the endorsement of a statement, once it is processed, is unaffected by cognitive load. Also note that it is possible that RRT effects increase given cognitive load. This is explainable due to the impact of cognitive load on the responding on incongruent trials (i.e., when the response rule is incongruent to the personal belief). It is likely that responding on such trials is influenced by cognitive load, in that it increases the task difficulty. This, however, also does not directly concern the process in question. Therefore, if RRT effects do not decrease in size given cognitive load, the endorsement of statements can be regarded as being automatic in the sense that they are independent on the availability of processing resources.

Third, when considering uncontrollability, the influence of two types of goals has to be considered, the goal to engage in the process, and the goal to stop, alter, or avoid the processing. With respect to the latter, studies on faking implicit measures are most informative (e.g., Degner, 2009; Klauer & Teige-Mocigemba, 2007; Teige-Mocigemba & Klauer, 2013). In such studies it is tested whether implicit measures are affected by the participants' knowledge about how the tasks work. Similar to the study of McKenna, Barnes-Holmes, and Barnes-Holmes (2007), who investigated faking IRAP effects, participants in a RRT faking study could be instructed to either cheat on the RRT, or, more specifically, to respond slower on congruent than on incongruent test blocks. If RRT effects are unaffected by such faking instructions, their occurrence can be considered independent of the goal to stop, alter, or avoid the personal endorsement of the statements. Interestingly, McKenna et al. (2007) found that IRAP effects were unaffected by faking instructions. Specifically, they asked three groups of participants to complete an IRAP task two times. After the first completion, each group of participants was instructed differently. One group was instructed to fake by responding slower on congruent test blocks. A second group was simply told to cheat on the task, without providing any information on how to cheat. The control group was not given any instructions. Crucially, no differences in the outcome were found between groups, indicating that participants were unable to cheat. However, whether the same is true for the RRT has to be tested empirically. Also, consider the second type of goal mentioned above, namely to engage in the personal endorsement of the stimuli of the RRT. Crucially, participants are not asked to express their personal beliefs. They are, however, asked to act as if they endorse a certain belief in one block of trials, and to act as if they do not endorse that belief in a second block of trials. Consequentially, in one of the two task-conditions, personal belief and instructed belief coincide. It is therefore hard to exclude the possibility that participants recode the task and simply respond in accordance with their personal belief on those trials on which instructed and personal beliefs coincide. It is furthermore imaginable that participants actively responded in contrast to their personal belief on trials on which their personal belief differed from the instructed task rule. Hence, at present, it is difficult to make strong claims about whether or not the endorsement of propositions is automatic in the sense of being valid in the absence of the intention to express the belief. Future work could examine this issue by determining the strategies that participants use while completing an RRT. Also, to investigate whether the goal to endorse statements is necessary for the occurrence of RRT effects, one could deploy a task in which participants are not required to endorse the statements in any way as part of the task rules. Consider, for instance, a task in which participants are presented with grammatically correct and incorrect statements that are either believed to be true or false. The words ‘true’ or ‘false’ are presented above and below the statements (e.g., Debey, De Houwer, & Verschuere, 2014). Participants are asked to indicate whether or not the statement is grammatically correct. Imagine that the categorization of personally endorsed statements is faster when these are flanked by the word true as compared to when they are flanked by the word false, and vice versa for personally rejected statements. The finding of such a systematic impact of the irrelevant flanker would be explainable by the conceptual overlap of the endorsement/rejection of the statement and the meaning of the flanker. The finding would therefore indicate that participants automatically endorse or reject the statements while processing them, even if they an endorsement of statements (in any way) is not demanded by the task.

Last, being unconscious of a process comprises of being unaware of the input (i.e., the statements), the output (i.e., the personal belief), or the process itself (i.e., the automatic endorsement of the personal belief and its influence on target categorizations). It seems obvious that participants are aware of the experimental stimuli, as they are both task-relevant and clearly visible. Moreover, oral feedback of participants who completed the studies presented in Chapters 4 and 5 revealed that participants can also be aware of the to-be-measured belief (i.e., racist beliefs and body image). Also, multiple students stated that responding in one condition was harder than in the other condition, based on their personal beliefs. However, a systematic investigation of these assumptions could shed more light on this issue.

On the basis of the above discussion of the nature of the endorsement of propositional statements and the conditions under which it can operate, it can be concluded that the endorsement of propositions is a propositional process which operates automatically at least with regard to one feature of automaticity. The observation of meaningful RRT effects in the presented experiments implies that propositions can be endorsed in a fast and, potentially, unintentional manner.[2] The findings substantiated the claim that the RRT can be used as tools to capture beliefs at the implicit level. As pointed out, more research is necessary to strengthen this claim with regard to other features of automaticity.

Conclusion

As discussed in detail above, the observations of automatic relational processing and automatic endorsement of propositions have to be regarded as instances of automatic propositional processing. The findings presented in this dissertation are difficult to reconcile with the central assumption of dual-process models of cognition that all

[2] Note, however, that this does not mean that the automatic endorsement of propositional knowledge is restricted to occur under this set of automaticity features. This will be discussed in detail in a subsequent section of this Chapter.

automatic processing must rely on associative processes (e.g., Rydell & McConnell, 2006). In light of the presented findings, the idea of two distinct modes of processing that are strictly separated along the lines of automaticity has to be discarded in favor of the idea that propositional processes can occur under conditions of automaticity. Nevertheless, it seems intuitively plausible that the presence or absence of different automaticity conditions can impact the degree of complexity of propositional processing. This idea is elaborated upon in the next section. The chapter will be concluded with a discussion of theoretical and practical implications and open questions of automatic propositional processing.

IMPLICATIONS AND OPEN QUESTIONS

On the complexity of automatic propositional processing

The assumption of dual-process models of cognition that automatic stimulus processing must be driven by associative processes limited the scope of automatic processing dramatically. As soon as a process was seen to be capable of occurring automatically, it had to be explained in associative terms (e.g., the processing of the valence of a stimulus as the automatic spread of activation from the stimulus concept to a valence node, see Fazio, 2001). The findings presented in this dissertation show, however, that this structural restriction is no longer tenable. As argued above, it seems sensible to consider the complexity of propositional processing to be determined by the automaticity conditions under which it operates. This assumption will be discussed with regard to the two types of automatic propositional processing investigated in this dissertation, i.e., automatic relational processing and automatic endorsement of propositions.

Automatic relational processing in the sequential priming task can be seen as an automatic construction of propositional knowledge about two concepts (i.e., the prime object and the reference object) and their relation (i.e., bigger or smaller) in the way that is consistent with the belief about how these concepts are related (e.g., the prime object ‘bike’ is smaller than the reference object ‘car’). Does the finding of relational priming effects in Chapter 2 allow for the conclusion that all types of comparative judgments regarding features of a prime stimulus and a reference stimulus can occur under the conditions of automaticity determined by the priming task? To answer this question, consider a number of priming studies conducted during the course of my PhD research that did not yield publishable results (see Appendix A, Experiments 4 to 6). The relational priming effects obtained in the size-judgment priming tasks of Chapter 2 allowed for inferences about the participants' beliefs about the relative size of physical objects. Experiments 4 to 6 of Appendix A were conducted to see whether it is possible to also assess psychologically more relevant beliefs in this way. In Experiment 5, I focused on beliefs about how people saw themselves (actual-self) and how they desired to be (ideal-self). Prime and target set comprised of adjectives of personality attributes (e.g., responsible, selfish, lazy, funny). In one task, participants were asked to categorize targets as being self-descriptive or not (i.e., actual-self priming task). In the second task, participants were asked to categorize targets as being desirable or not (ideal-self priming task). Prime stimuli were expected to influence target categorization based on their self-descriptiveness and desirability in the two tasks, respectively. Within the actual-self priming task, prime adjectives that were descriptive of the participant were expected to facilitate responses towards self-descriptive target adjectives as compared to non-descriptive target adjectives. Non-descriptive prime adjectives in contrast were expected to facilitate responses towards non-descriptive target adjectives as compared to self-descriptive target adjectives. Desirable and non-desirable prime stimuli were expected to affect the judgments in the ideal-self priming task in the same way. Crucially, for each individual participant, the prime set was chosen to contain adjectives that were descriptive but undesired (e.g., lazy) or desired but not descriptive (e.g., funny). Similar to the medium-sized objects that were used as primes in the experiments of Chapter 2, these adjectives were expected to influence target categorization differently between tasks. For example, imagine a person that rated the attribute 'responsible' as highly descriptive and desirable, 'selfish' as both non-descriptive and undesirable, and 'lazy' as descriptive but undesirable. When categorizing targets on the basis of self-descriptiveness, the prime 'lazy' was expected to facilitate responses towards the descriptive target adjective 'responsible' as compared to non-descriptive target adjective 'selfish'. In contrast, when targets were categorized on the basis of desirability, the same prime ‘lazy’ was expected to facilitate the categorization of the undesired target adjective ‘selfish’ as compared to the desired target adjective ‘responsible’. Priming effects would therefore be indicative of the participants’ beliefs regarding their actual and ideal self. However, no such priming effects were observed in Experiment 5, or in the similar Experiments 4 and 6.

The absence of priming effects in these experiments indicates that processing prime stimuli in accordance with actual- and ideal-self beliefs did not result in the construction of propositional knowledge. However, given the same conditions of automaticity, the processing of prime stimuli in accordance with beliefs about their relative size resulted in the construction of propositional knowledge. It therefore has to be argued that whether stimuli are relationally processed is dependent on the type of comparative judgment and the task-relevant features of a prime stimulus and a reference stimulus. This intuitively makes sense if one considers the two types of judgments that were focused upon. The size of physical objects is factual independent of the context, whereas judgments about the descriptiveness or desirability of a personality attribute are clearly context-dependent. For instance, whether someone considers him or herself as lazy depends on the knowledge that is used to make that judgment. Even the most lazy person can think of instances in which the opposite is true (i.e., in which he/she is not lazy; e.g., with respect to regular physical activity). The ease of a judgment relies on the knowledge that is called upon. If that knowledge is ambiguous, the judgment should be harder as compared to when it is unambiguous (e.g., judgments of size).

The findings in Chapters 4 and 5 indicate that the RRT, in contrast to the sequential priming task, allows for the automatic processing of stimuli with regard to personal, potentially ambiguous, beliefs. One could argue that this is due to the fact that processing in the RRT and the priming task differ with regard to the degree of processing automaticity. As discussed above, propositional processing in the priming task operates under at least three features of automaticity, whereas propositional processing in the RRT can be considered automatic only with respect to speed and, potentially, intentionality. Such reasoning would, however, disregard the fact that the automatic propositional processing differs between the two tasks. As indicated above, automatic processing in the priming tasks consists of the construction of propositional knowledge in line with the personal belief. Automatic processing in the RRT is, in contrast, restricted to the evaluation of propositions as being true or false (i.e., endorsement or rejection of propositions). The discrepancy in the findings in the size-judgment task of Chapter 2 and the actual- and ideal-self judgment tasks of Appendix A therefore only indicates that the construction of propositional knowledge is not unconditionally possible given the set of automaticity features uphold by the specific priming task, but, as agued above, instead depends on the specific judgment and stimulus material. It cannot, however, be argued that an automatic endorsement of ambiguous propositional knowledge can only operate under the less stringent automaticity conditions uphold by the RRT. Rather, it has to be tested whether the automatic processing that occurs in the RRT can also operate given stricter conditions of automaticity. Specifically, it has to be tested whether the automatic endorsement of statements is also possible given the automaticity conditions of the priming paradigm. Imagine, for instance, a priming task, in which the stimuli “I want to be” or “I am” are presented as cue on trial offset. The cue is followed by personality attributes as prime and target words (e.g., lazy, funny). Participants are asked to categorize targets with regard to the cues as being true or false. Crucially, in such a task, one can assume that the stimulus compounds (i.e., cue and target, and cue and prime, respectively) can be processed as readily constructed propositions. Given the unintentional misapplication of the categorization task to be cue-prime-compound, an automatic endorsement or rejection of the statement is imaginable. If this is observed, that is, if target categorization is systematically influenced by the personal beliefs about the truthfulness of the proposition formed by prime and cue, it would suggest that the endorsement of propositions can also occur under the stricter set of automaticity conditions of the priming task.

To summarize, the findings that automatic processes can be propositional in nature led to the question how the complexity of a propositional process should be best defined. With regard to the two processes investigated in this dissertation, it was argued that a critical assessment of the complexity of a process is possible only by reflecting the set of automaticity features under which that process can operate. These seem to be determined by the properties of the specific experimental procedure that is used to investigate the process (e.g., stimulus set, task, time-course). Crucially, knowing under which automaticity features a process can or cannot operate is important when hypothesizing about its influence on behavior. Implicit measures are regarded as laboratory equivalents of automatic stimulus processing in real life. Predictions about real-life behavior with implicit measures therefore heavily rely on knowing of the automaticity conditions under which specific processes occur in a specific real-life setting to be capable of deploying an implicit measure that fits these automaticity conditions.

Propositional vs. associative implicit measures

Traditional associative implicit measures, like the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) and the evaluative priming task (Fazio, Jackson, Dunton, & Williams, 1995) were designed to capture the extent to which certain concepts are related in memory (Hughes, Barnes-Holmes, & Vahey, 2012). They can be used to capture the extent to which a certain class of stimuli (e.g., blacks) is related with a positive or negative valence. Crucially, however, associative implicit measures do not generally restrict the way in which stimuli are related to one another. Propositional implicit measures, like the RRT and the IRAP, in contrast restrict stimulus processing. Consequentially, these measures always probe specific cognitions rather than an unspecific relatedness. This is of importance, as in many cases the precise way in which concepts are processed is of crucial interest.

Consider, for instance, the findings of De Raedt, Schacht, Franck, and De Houwer (2006). They observed that a group of depressed patients and a non-depressed control group did not differ in their scores of implicit self-esteem as measured with a standard self-esteem IAT. However, explicit measures clearly indicated that depressed patients possessed a significantly lower level of self-esteem as compared to the non-depressed control group. Remue, De Houwer, Barnes-Holmes, Vanderhasselt, and De Raedt (2013; see also Remue, Hughes, De Houwer, & De Raedt, 2014) provided a lucid explanation for this puzzling finding by arguing that depressed people differ from non-depressed people in the way in which they automatically process information. Using two IRAPs, they showed that dysphoric students had lower scores on an IRAP designed to capture the belief “I am good” than on an IRAP designed to capture the belief “I want to be good” whereas non-dysphoric students showed the reversed pattern. Remue et al. (2013) argued that the positive scores of the self-esteem IAT might reflect different types of processing of the stimuli for the two groups of participants. They proposed that positive scores of the self-esteem IAT might reflect processing of the attitude stimuli in line with their positive actual self-esteem for non-dysphoric students, whereas the same positive score reflected stimulus processing in line with a positive ideal self-esteem for dysphoric participants.

The work by Remue et al. (2013) underscore the benefits of using implicit measures that are sensitive to how concepts are related to another. This need for a higher specificity in implicit measures is also reflected in the development of multiple “propositionalized” variants of the IAT. In the personalized IAT (pIAT) of Olson and Fazio (2004; see also Dewitte & De Houwer, 2008), participants are asked to classify attribute stimuli as instances of the categories “I like” or “I dislike” instead of as instances of the categories “pleasant” or “unpleasant” used in the standard IAT of Greenwald et al. (1998). As argued by Olson and Fazio (2004) IAT scores reflected the participants’ personal rather than normative, societally shared evaluations. Yoshida, Peach, Zanna, and Spencer (2012), in contrast, used the attribute category labels “most people like” and “most people do not like” to focus stimulus processing to societally shared knowledge rather than personal evaluations. In each case, the reasoning behind the change in categories of the IAT was driven the assumption that specific knowledge (e.g., the personal rather than the societal evaluation of an out-group) is related to specific behavior (e.g., discriminative behavior against an out-group). Crucially, as these instances of the IAT require the processing of stimuli in a specific relation (e.g., with regard to personal beliefs), they can be seen as propositional implicit measures.

To summarize, implicit measures are used by psychologists who are interested in measuring cognitions (e.g., attitudes, beliefs) to understand, control, and predict human behavior. The more stringently an implicit measure restricts stimulus processing to specific relational information, the more precise the measurement outcome should reflect a specific cognition. The prediction of automatic behavior, being based on specific cognitions, should therefore be predicted best with propositional implicit measures that restrict stimulus processing accordingly. If, for instance, racist behavior is mainly guided by personal beliefs about an out-group, the implicit measurement procedure that assesses these specific beliefs should best predict automatic racist behavior. If stimulus processing is not restricted, the interpretation of implicit measures is prone to unwarranted assumptions about how participants process the experimental stimuli. As these assumptions could be wrong, such implicit measures should in general be less predictive of behavior than propositional implicit measures that restrict stimulus processing to mirror stimulus processing in real life.

Further benefits and potential restrictions of propositional implicit measures

Since their introduction two decades ago, implicit measures have gained enormous popularity as a means to indirectly assess cognitions and predict behavior. As argued by Fazio and Olson (2003), implicit measures are suited to predict behavior particularly in situations in which people are not motivated or capable of controlling the impact of their cognitions on behavior. Implicit measures have therefore been used intensively in clinical addiction research (e.g., Stacy & Wiers, 2010; Teachman, Cody, & Clerkin, 2010; Wiers, Houben, Roefs, De Jong, Peter, & Etchison, 2010; Wiers, Van Woerden, Smulders, & De Jong, 2002), and research on racial stereotypes and racist behavior (e.g., Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio et al., 1995; McConnell & Leibold, 2001). In such instances, implicit and explicit measures can be expected to diverge, as explicit measures are likely to be influenced by the motivation to correct responses to match social acceptance (e.g., Dunton & Fazio, 1997; Greenwald & Nosek, 2008; Petty, Fazio, & Briñol, 2008). Accordingly, the use of implicit measures should be most beneficial in such instances. When people, in contrast, are willing and capable of reporting their beliefs, implicit and explicit measures converge. The finding that both automatic and non-automatic processing relies on the same propositional representations implies that under these circumstances, propositional implicit measures should highly correlate with explicit measures of the same belief. This allows to test the accuracy of implicit measures.

Consider, for instance, the assessment of actual and ideal body image with the RRT in Chapter 5. It was found the RRT scores of actual and ideal body image, but also the aggregated scores of the explicit ratings of the RRT statements, were valid indicators of the degree to which participants were satisfied with their body. Such a finding was expected, as students in the anonymous setting of an experiment were expected to report their true beliefs in the explicit ratings. In other words, the fact that implicit and explicit measures predicted body dissatisfaction speaks for the accuracy of the RRTs. However, as argued in Chapters 3 and 5, there are good reasons to suspect that in applied research contexts, implicit and explicit measures of body image converge. Patients that suffer from eating disorder might for instance be inclined to fake on explicit measures of ideal and actual body image because of therapeutic consequences like compulsory admission. In such settings, explicit measures should consequentially be less informative than implicit ones that focus on the same beliefs.

A remaining question with regard to the versatility of the RRT is whether it can also be used to measure multiple beliefs at once. Questionnaires allow assessing multiple facets of a belief, and even qualitatively different beliefs. Each item of a questionnaire is rated independently on the level of agreement with that statements or question. The RRTs presented in Chapters 4 and 5, in contrast, assessed only one specific belief each. However, it is easily imaginable to adjust the RRT to assess multiple facets of a belief or even qualitatively different beliefs. For that, target statements have to be used that reflect these different beliefs, and the response rules have to be adjusted to apply to all beliefs under consideration. Consider, for instance, the assessment of alcohol-related expectancies, i.e., beliefs that alcohol consumption will serve a certain function or result in a specific desirable consequence. The Alcohol Expectancy Questionnaire (AEQ; Brown, Christiansen, & Goldman, 1987) focuses on multiple such expectancies at once (e.g., general positive or negative affect, relaxation and tension reduction, cognitive and motor impairment, sexual enhancement; for a more comprehensive list of factors, see Brown et al., 1987, Chart 1). Sample items of the factor *sexual enhancement* are, for instance, “alcohol makes sexual experiences easier and more enjoyable” and “I often feel sexier after I've had a few drinks”. Sample items of the factor *relaxation and tension reduction* are, for instance, “alcohol decreases muscular tension” “alcohol makes me worry less“. Crucially, these items, or abbreviated versions of them, could be used as target statements in an alcohol-expectancy RRT that addresses multiple different types of alcohol-related expectancies. Crucially, the overarching response rules would have to be adapted to apply to all types of expectancies that are considered within the RRT (e.g., "respond as if you believe that alcohol has positive effects"). Separate RRT scores for each type of target statement (e.g., beliefs about sexual enhancement) can be calculated by using only the subset of trials on which statements regarding that belief were presented. The main restriction of the RRT in assessing multiple beliefs therefore seems to be the phrasing of response rules so that they apply to all beliefs under consideration. This should become increasingly difficult with the number of different beliefs included in one RRT.

To summarize, it was argued that based on the finding that propositions, rather than associations, underlie automatic processing, implicit measurement procedures should be capable of assessing the same beliefs as explicit measures. This, however, is only possible if stimulus processing in implicit measurement procedures is restricted. Otherwise, uncertainty remains about what specific belief was accessed. If a person lacks motivation or introspection, explicit measures become unreliable, whereas propositional implicit measures are expected to remain credible indicators of beliefs. Further, propositional implicit measurement procedures like the RRT are theoretically capable of assessing multiple beliefs within one task, as long as the response rules are phrased in a way that they are applicable to all beliefs under consideration.

FINAL CONCLUSION

The research presented and discussed in my dissertation provides supporting evidence for the assumption that, in line with single-process propositional accounts of cognition, human behavior is guided by propositional processes (De Houwer, 2009, 2014; Mitchell et al., 2009). Conversely, the reported findings seems difficult to reconcile (at least in a parsimonious manner) with classic dual-process models of cognition (for an overview, see De Houwer, 2014). In any event, the present studies clearly show that automatic stimulus processing can be much more complex than previously assumed.

REFERENCES

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, *32*, 169–177.

Brown, S. A., Christiansen, B. A., & Goldman, M. S. (1987). The Alcohol Expectancy Questionnaire: an instrument for the assessment of adolescent and adult alcohol expectancies. *Journal of Studies on Alcohol*, *48*, 483–491. doi: 10.15288/jsa.1987.48.483

Cash, T. F., & Szymanski, M. L. (1995). The Development and Validation of the Body-Image Ideals Questionnaire. *Journal of Personality Assessment*, *64*, 466–477. doi: 10.1207/s15327752jpa6403_6

Debey, E., De Houwer, J., & Verschuere, B. (2014). Lying relies on the truth. *Cognition*, *132*, 324–334. doi: 10.1016/j.cognition.2014.04.009

Degner, J. (2009). On the (un-) controllability of affective priming: Strategic manipulation is feasible but can possibly be prevented. *Cognition & Emotion*, *23*, 327–354. doi: 10.1080/02699930801993924

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*, 1–20. doi: 10.3758/LB.37.1.1

De Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass*, *8*, 342–353. doi: 10.1111/spc3.12111

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*, 347–368. doi: 10.1037/a0014211

De Raedt, R., Schacht, R., Franck, E., & De Houwer, J. (2006). Self-esteem and depression revisited: Implicit positive self-esteem in depressed patients? *Behaviour Research and Therapy*, *44*, 1017–1028. doi: 10.1016/j.brat.2005.08.003

Dewitte, M., & De Houwer, J. (2008). Proximity and distance goals in adult attachment. *European Journal of Personality*, *22*, 675–694. doi: 10.1002/per.696

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *33*, 510–540. doi: 10.1006/jesp.1997.1331

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin*, *23*, 316–326. doi: 10.1177/0146167297233009

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition & Emotion*, *15*, 115–141. doi: 10.1080/02699930125908

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013–1027. doi: 10.1037/0022-3514.69.6.1013

Fazio, R. H., & Olson, M. A. (2003). Implicit Measures in Social Cognition Research: Their Meaning and Use. *Annual Review of Psychology*, *54*, 297–327. doi: 10.1146/annurev.psych.54.101601.145225

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology*, *40*, 535–542. doi: 10.1016/j.jesp.2003.10.005

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Greenwald, A. G., & Nosek, B. A. (2008). Attitudinal dissociation: What does it mean. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 65–82). New York, NY: Psychology Press.

Hermans, D., Crombez, G., & Eelen, P. (2000). Automatic attitude activation and efficiency: The fourth horseman of automaticity. *Psychologica Belgica*.

Higgins, E. T. (1989). Self-Discrepancy Theory: What Patterns of Self-Beliefs Cause People to Suffer? *Advances in Experimental Social Psychology*, *22*, 93–136. doi: 10.1016/s0065-2601(08)60306-8

Hughes, S., Barnes-Holmes, D., & Vahey, N. (2012). Holding on to our functional roots when exploring new intellectual islands: A voyage through implicit cognition research. *Journal of Contextual Behavioral Science*, *1*, 17–38. doi: 10.1016/j.jcbs.2012.09.003

Klauer, K. C., & Teige-Mocigemba, S. (2007). Controllability and resource dependence in automatic evaluation. *Journal of Experimental Social Psychology*, *43*, 648–655. doi: 10.1016/j.jesp.2006.06.003

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, *37*, 435–442. doi: 10.1006/jesp.2000.1470

McKenna, I. M., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2007). Testing the fake-ability of the Implicit Relational Assessment Procedure (IRAP): The first study. *International Journal of Psychology and Psychological Therapy*, *7*, 253–268.

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*, 183–198. doi: 10.1017/S0140525X09000855

Moors, A. (2016). Automaticity: Componential, causal, and mechanistic explanations. *Annual Review of Psychology*, *67*.

Moors, A., & De Houwer, J. (2006). Automaticity: A Theoretical and Conceptual Analysis. *Psychological Bulletin*, *132*, 297–326. doi: 10.1037/0033-2909.132.2.297

Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extrapersonal associations on the Implicit Association Test: personalizing the IAT. *Journal of Personality and Social Psychology*, *86*, 653–667. doi: 10.1037/0022-3514.86.5.653

Petty, R. E., Fazio, R. H., & Briñol, P. (2008). The new implicit measures: An overview. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 3–18). New York, NY: Psychology Press.

Remue, J., De Houwer, J., Barnes-Holmes, D., Vanderhasselt, M.-A., & De Raedt, R. (2013). Self-esteem revisited: Performance on the implicit relational assessment procedure as a measure of self- versus ideal self-related cognitions in dysphoria. *Cognition & Emotion*, *27*, 1441–1449. doi: 10.1080/02699931.2013.786681

Remue, J., Hughes, S., De Houwer, J., & De Raedt, R. (2014). To Be or Want to Be: Disentangling the Role of Actual versus Ideal Self in Implicit Self-Esteem. *PLoS ONE*, *9*, e108837. doi: 10.1371/journal.pone.0108837

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: a systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*, 995–1008. doi: 10.1037/0022-3514.91.6.995

Stacy, A. W., & Wiers, R. W. (2010). Implicit cognition and addiction: a tool for explaining paradoxical behavior. *Annual Review of Clinical Psychology*, *6*, 551–575. doi: 10.1146/annurev.clinpsy.121208.131444

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247. doi: 10.1207/s15327957pspr0803_1

Strauman, T. J., Vookles, J., Berenstein, V., Chaiken, S., & Higgins, E. T. (1991). Self-discrepancies and vulnerability to body dissatisfaction and disordered eating. *Journal of Personality and Social Psychology*, *61*, 946–956. doi: 10.1037/0022-3514.61.6.946

Teachman, B. A., Cody, M. W., & Clerkin, E. M. (2010). Clinical applications of implicit social cognition theories and methods. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 489–564). New York, NY: Guilford Press.

Teige-Mocigemba, S., & Klauer, K. C. (2013). On the controllability of evaluative-priming effects: Some limits that are none. *Cognition & Emotion*, *27*, 632–657. doi: 10.1080/02699931.2012.732041

Wiers, R. W., Houben, K., Roefs, A., De Jong, Peter, & Etchison, S. (2010). Implicit cognition in health psychology: why common sense goes out the window. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 463–488). New York, NY: Guilford Press.

Wiers, R. W., Van Woerden, N., Smulders, F. T. Y., & De Jong, P. J. (2002). Implicit and explicit alcohol-related cognitions in heavy and light drinkers. *Journal of Abnormal Psychology*, *111*, 648–658. doi: 10.1037/0021-843X.111.4.648

Williamson, D. A., Gleaves, D. H., Watkins, P. C., & Schlundt, D. G. (1993). Validation of self-ideal body size discrepancy as a measure of body dissatisfaction. *Journal of Psychopathology and Behavioral Assessment*, *15*, 57–68. doi: 10.1007/BF00964324

Yoshida, E., Peach, J. M., Zanna, M. P., & Spencer, S. J. (2012). Not all automatic associations are created equal: How implicit normative evaluations are distinct from implicit attitudes and uniquely predict meaningful behavior. *Journal of Experimental Social Psychology*, *48*, 694–706. doi: 10.1016/j.jesp.2011.09.013

APPENDIX A

ADDITIONAL RESEARCH

OVERVIEW

Research that was conducted during the course of my PhD that did not yield publishable results was not included in the main body of this thesis. However, in order for this dissertation to be a comprehensive reflection of my work, these experiments are briefly summarized in the following sections. Experiments 1 and 2 constitute the first two instances in which task-misapplication was investigated, the mechanism by which automatic (relational) stimulus processing is thought to take place in the sequential priming paradigm (cp. Chapter 2). Experiment 3 constitutes an unsuccessful attempt to replicate the observation of relational priming effects of Chapter 2 with different task and stimulus material. Experiments 4 to 6 are three priming experiments that focused on the implicit measurement of self-relevant beliefs. Experiment 7 constitutes a successful proof of principle experiment of the Relational Responding Task preceding the experiments presented in Chapters 4 and 5.

TASK MISAPPLICATION

In Chapter 2 the mechanism of task misapplication (e.g., Dehaene et al., 1998; Van Opstal, Gevers, Osman, & Verguts, 2010) was investigated as a means by which priming effects can be explained as resulting from the automatic construction of propositional knowledge about the prime stimuli rather than the activation of knowledge associated to the prime stimuli. The mechanism of task misapplication essentially constitutes of the unintended misapplication of the target task to the primes, resulting in the processing of the primes in line with the target task.

Experiment 1

In a first priming experiment, stimuli were words that differed on two semantic features, namely their evaluative tone (i.e., being either positive or negative) and their liveliness (i.e., denoting either a person or an object). The stimulus material (e.g., friend, thief, present, gun) was taken from the study of Spruyt, De Houwer, and Hermans (2009). Prime and target stimuli were presented in blue or orange color on different trials of the experiment. Seventy-two participants were asked to respond to the targets on the basis of their liveliness when they were presented in blue. When targets where presented in orange, participants were asked to respond to the targets on the basis of their evaluative valence. The color of the prime of each trial either matched or mismatched the color of the target.

Based on the assumption that participants would automatically misapply the target task to the primes, a semantic priming effect was expected on trials on which prime and target were both presented in blue. Similarly, an evaluative priming effect was expected on trials on which prime and target were both presented in orange. In contrast, on trials of a color mismatch between prime and target no priming effects were expected to arise due to the absence of a pattern between the responses indicated by primes and targets (i.e., as primes should be processed on the basis of their liveliness when targets had to be categorized on the basis of their evaluative tone, or vice versa). However, neither a semantic priming effect was found on trials on which stimuli were presented in blue, $F < 1$, nor a evaluative priming effect on trials on which stimuli were presented in orange, $F < 1$.

Experiment 2

One possible explanation for the absence of priming effects in Experiment 1 was that participants did not process primes in accordance with any task, as the target task was indicated by the color of the target, and thus known to the participant only after the presentation of the prime. Consequentially, task misapplication at the point of prime presentation was unlikely. Moreover, response latencies to the targets were rather long (i.e., about 1000 ms), possibly indicating that participants only at the time of the target presentation chose which task to apply. To increase the likelihood of a misapplication of the target task to the primes, in this study the point in time at which participants were informed about the relevant task on each trial was antedated to the trial onset. Using the same stimulus material as in the previous experiment, 37 participants were asked to categorize targets on either their feature of evaluative tone or liveliness depending on a cue that was presented on trial onset. However, no priming effects were obtained with this setup either. Neither a significant semantic priming effect was found on trials on which stimuli were presented in blue, $F < 1$, nor was a significant evaluative priming effect found on trials on which stimuli were presented in orange, $F < 1$.

RELATIVE EVALUATIVE JUDGMENTS

Experiment 3

Analogous to the size-judgment priming task, participants in this experiment (N = 46) were asked to categorize words (i.e., names of animals; taken from Moors et al., 2013) as positive or negative in relation to one of two reference animals presented on trial onset. Stimuli were chosen based on rating norms as belonging to one of the three groups ‘positive’ ‘neutral’ and ‘negative’). The two reference animals were slightly positive (but to a lesser extent than the stimuli of the positive group; i.e., a swan), and slightly negative (but to a lesser extent than the stimuli of the negative group; i.e., a mosquito). On trials on which targets were compared to the slightly negative reference object mosquito, primes of the neutral stimulus group were expected to facilitate the categorization of positive targets as compared to negative targets. When targets were compared to the slightly positive reference object swan, neutral primes were expected to facilitate the categorization of negative targets as compared to positive targets. However, on critical trials (i.e., when primes belonged to the group of neutrally valenced animals), the crucial interaction indicating a relational priming effect was not significant, $F < 1$.

ASSESSING IMPLICIT BELIEFS WITH THE PRIMING TASK

In Chapter 3 I presented results obtained with two Implicit Relational Assessment Procedures (IRAP; Barnes-Holmes et al., 2006). Specifically, it was demonstrated that beliefs about actual and ideal body image can be assessed. Crucially, beliefs reflecting both constructs differed in how the inherent concepts *self* and *body size* were related to each other. The actual body image can be represented by beliefs that relate self and body size descriptively (e.g., I *am* thin). Beliefs reflecting the ideal body image, in contrasts, are characterized by a relation of desirability (e.g., I *want to be* thin). In following three experiments the descriptive and desirability relation were assessed using the sequential priming paradigm.

Experiment 4

In a first experiment, 34 participants were asked to categorize target words as being self-descriptive (response options: ‘am’ and ‘am not’) or desirable for them (response options: ‘want to be’ and ‘don’t want to be’) in two successive priming tasks. Primes were adjectives (e.g., responsible, reliable, intelligent, heartless, stupid, selfish, lazy, funny) and chosen individually based on explicit ratings of descriptiveness and desirability prior to the priming task. Targets in the descriptive priming task were five words that were descriptive (i.e., student, Belgian, female, young) and non-descriptive (i.e., lawyer, Dutch, male, old) of the participant, respectively. Accordingly, targets in the desirability priming task were five words that were desirable (i.e., psychologist, rich, healthy, respected) or undesirable (i.e., Soldier, dead, sick, lonely) for the participant.

Different patterns of priming effects were expected in the two priming tasks. Crucially, primes that differed in their descriptiveness and desirability should facilitate agreeing responses (i.e., ‘am’ or ‘want to be’) as compared to denying responses (i.e., ‘am not’ or ‘don’t want to be’) in one task, and vice versa in the other task. For instance, the self-descriptive but undesirable adjective ‘lazy’ was expected to facilitate the categorization of descriptive target words (e.g., female) relative to non-descriptive target words (e.g., male) in the descriptive priming task, and to facilitate the categorization of undesirable target words (e.g., sick) relative to desirable target words (e.g., rich) in the desirable priming task. Such a finding would have indicated that primes can be processed in accordance with the (misapplied) target categorization task.

However, no significant priming effects were found within the tasks, $F < 1$, for the interaction of the descriptiveness of primes and targets in the descriptive priming task, and $F(1, 27) = 1.45$, $p = .238$, $\eta\rho^2 = .05$, for the desirability of primes and targets in the desirability priming task.

Experiment 5

This experiment was conducted to exclude the possibility that the absence of priming effects was due to the fact that prime and target set of in the previous experiment differed too strongly to allow for a misapplication of the task. Therefore, in this experiment both the prime and the target set exclusively contained attitude adjectives (e.g., responsible, reliable, intelligent, heartless, stupid, selfish, lazy, funny). These were again chosen individually for each participant based on prior explicit ratings of their descriptiveness and desirability. Identical to the previous experiment, participants (N = 43) were asked to categorize target words as being self-descriptive (response options: ‘am’ and ‘am not’) or desirable (response options: ‘want to be’ and ‘don’t want to be’) in the two consecutive priming tasks. However, again no significant priming effects were found. Neither the interaction of prime and target descriptiveness in the descriptive priming task, nor the interaction of prime and target desirability in the desirability priming task yielded significant results, $F < 1$, and $F(1, 41) = 3.34$, $p = .075$, $\eta\rho^2 = .08$, respectively. Moreover, explicit ratings of the descriptiveness and desirability of the primes taken after the priming task did not correspond well with the ratings of the stimuli taken prior to the priming task. This finding suggested that participants were unable to consistently rate adjectives as descriptive and desirable.

Experiment 6

Thus, in our third attempt, I restrained from manipulating the stimulus material in order to receive stimuli that differed in their relation of descriptiveness and desirability. Instead I relied on groups of people that possessed specific but differing beliefs about related constructs, namely beliefs about actual and ideal body image. After the successful dissociation of these beliefs with the IRAP (Chapter 3), I aimed at capturing them with two priming tasks focusing on the descriptive and desirable relation between self and body size.

In two priming tasks, prime words were exemplars of the two categories *thin* and *overweight*. Target words of the actual body image priming task were descriptive and non-descriptive of the participants (i.e., ‘student’, ‘female’, ‘young’, ‘lawyer’, ‘male’, ‘old’). Targets of the ideal body image priming task descriptive of the participants desired or undesired future selves (i.e., ‘successful’, ‘rich’, ‘psychologist’, ‘sick’, ‘poor’, ‘soldier’). Sixty-three body satisfied and body dissatisfied participants were asked to categorize targets as being descriptive or desirable of their selves in two priming tasks. Prime words in both tasks were expected to influence responses towards the target stimuli depending on the participants’ implicit beliefs about their actual and ideal body image. Specifically, for participants that believed to be thin (i.e., that possessed a thin actual body image) I expected primes of the category *thin* to facilitate the categorization of descriptive target words relative to non-descriptive target words. For participants that believed not to be thin (i.e., that possessed an overweight actual body image) I expected primes of the category *overweight* to facilitate the categorization of descriptive target words relative to non-descriptive target words. Moreover, for these participants I expected a reversed pattern in the ideal body image priming task.

Specifically, for participants that desired to be thin (i.e., possessing a thin ideal body image) I expected primes of the category thin to facilitate the categorization of desired target words relative to the categorization of undesired target words.

A significant overall priming effect of 8 ms was found in the desirability priming task, $F(1, 60) = 4.74$, $p = .033$, $\eta\rho^2 = .07$, indicating faster responses to desirable target words when preceded by primes of the category thin as compared to primes of the category overweight, and faster responses to non-desirable target words when preceded by primes of the category overweight as compared to primes of the category thin. This interaction, however, did not differ significantly between participants with high and low body dissatisfaction, $F(1, 59) = 2.45$, $p = .123$, $\eta\rho^2 = .04$. No priming effects were found for the descriptive priming task, $F < 1$.

Despite the fact that people high and low in body dissatisfaction differ in their believes about their actual and ideal body image (cp. Chapter 3), I was unable to capture these differences with the priming task. This led to the assumption that the priming task as an irrelevant feature task might have been not sensitive enough for the measurement of implicit beliefs about the self (cp. Discussion in Chapter 6).

PROOF OF PRINCIPLE OF THE RELATIONAL RESPONDING TASK

Experiment 7

Prior to the successful demonstration of the Relational Responding Task (RRT) in Chapter 4, a proof of principle experiment of the RRT with 65 participants was conducted, using statements that contained a relation of size (i.e., smaller than, larger than) and were either objectively true or false (e.g., ‘an apple is smaller than a boat’, ‘a banana is larger than a car’). Scores of the RRT were expected to indicate an implicit belief of objectively true statements and a disbelief of objectively false statements. Furthermore, it was tested whether scores of the RRT were independent of the order of consistent blocks (i.e., responding in line with the objective truth of statements) and inconsistent blocks (i.e., responding in contrast to the the objective truth of statements). Reassuringly, no effects of order was found, all $F < 1$. Most crucially, the average RRT score of 0.25 was sign. larger than zero, $t(64) = 4.41$, $p < .001$, indicating faster responses to the statements in line with the rule that stated objective truth (i.e., “respond as if you believe that fruits are smaller than vehicles”).

REFERENCES

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, *32*, 169–177.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., … Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature*, *395*, 597–600. doi: 10.1038/26967

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A.-L., … Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods*, *45*, 169–177. doi: 10.3758/s13428-012-0243-8

Spruyt, A., Houwer, J. D., & Hermans, D. (2009). Modulation of automatic semantic priming by feature-specific attention allocation. *Journal of Memory and Language*, *61*, 37–54. doi: 10.1016/j.jml.2009.03.004

Van Opstal, F., Gevers, W., Osman, M., & Verguts, T. (2010). Unconscious task application. *Consciousness and Cognition*, *19*, 999–1006. doi: 10.1016/j.concog.2010.05.002

APPENDIX B

NEDERLANDSTALIGE SAMENVATTING

Het onderzoek dat in deze dissertatie wordt voorgesteld is gebaseerd op de assumptie dat propositionele processen automatisch gedrag (kunnen) aansturen, een hypothese die centraal staat binnen zgn. propositionele één-proces modellen van cognitie (De Houwer, 2009, 2014; Mitchell, De Houwer, & Lovibond, 2009). Deze modellen contrasteren sterk met (de meer gangbare) twee-proces modellen van cognitie die vooropstellen dat automatisch gedrag wordt aangestuurd door associatieve processen (bv., Gawronski & Bodenhausen, 2006; Strack & Deutsch, 2004). Twee-proces modellen veronderstellen twee manieren van informatieverwerking die respectievelijk verantwoordelijk zijn voor automatisch en niet-automatisch gedrag. Automatisch gedrag zou aangestuurd worden door associatieve processen, i.e.., de automatische activatie van associaties tussen concepten binnen een associatief netwerk. Niet-automatisch gedrag, daarentegen, wordt verondersteld te worden aangestuurd door propositionele processen. De complexiteit van informatie in het propositioneel systeem verschilt van de informatie opgeslagen in het associatieve systeem. In tegenstelling tot associaties omvatten proposities gekwalificeerde relaties tussen concepten die gekenmerkt worden door een (subjectief of objectief) waarheidsgehalte. In Hoofdstuk 2 onderzoek ik het idee dat propositionele representaties (i.e., informationele eenheden die informatie bevatten over hoe concepten gerelateerd zijn; zie De Houwer, 2009, 2014) kunnen ontstaan onder automaticiteitscondities. In vier sequentiële primingstudies (bv., Fazio, Sanbonmatsu, Powell, & Kardes, 1986) werden deelnemers gevraagd om de fysieke grootte van een reeks targetobjecten te beoordelen, relatief ten opzichte van een eerder klein (i.e., een voetbal) of eerder groot (i.e., een auto) referentieobject. De set van cruciale primes bestond uit objecten die groter waren dan de kleine referentieobjecten maar kleiner dan de grote referentieobjecten (bv., een fiets). Resultaten toonden aan dat de impact van deze cruciale primes op de snelheid waarmee een target kon gecategoriseerd worden afhankelijk was van de grootte van het referentieobject. Deze interactie, ook wel het relationeel primingeffect genoemd, toont aan dat deelnemers de grootte van de primes verwerkten in relatie tot de grootte van het referentieobject. Aangezien dit effect tot stand kwam onder automaticiteitscondities bieden deze studies bovendien steun voor de idee dat propositionele informatieverwerking kan plaatsvinden onder automaticiteitscondities. Een tweede onderzoekslijn was gebaseerd op de hypothese dat de verwerking van propositionele informatie hand in hand gaat met de automatische evaluatie van deze informatie als waar of vals. Indien deze hypothese juist is, dan zou het in principe mogelijk moeten zijn om een indirecte meetprocedure te ontwikkelen die toelaat om indirecte overtuigingen (‘beliefs’) te capteren (i.e., proposities die spontaan geëvalueerd worden als waar door een bepaald individu). Het psychologisch construct van *lichaamsontevredenheid* heeft een belangrijke rol gespeeld bij het testen van deze assumptie. Lichaamsontevredenheid kan gezien worden als een negatieve attitude ten opzichte van het eigen lichaam die ontstaat wanneer men een discrepantie ervaart tussen het eigenlijke fysieke voorkomen (d.i., het *actuele* lichaamsbeeld) en geïnternaliseerde idealen over fysieke aantrekkelijkheid (d.i., het ideale lichaamsbeeld, b.v., Cash, 1989). Beide constructen, het actuele en het ideale lichaamsbeeld, zijn zeer gelijkend met uitzondering van de manier waarop de concepten *zelf* en *lichaamsgrootte* gerelateerd zijn. Proposities over het actuele lichaamsbeeld worden gedefinieerd door een beschrijvende relatie tussen het zelf en lichaamsgrootte (bv., ik *ben* dun). Proposities over het ideale lichaamsbeeld worden gedefinieerd in termen van wenselijkheid (bv. Ik *wil* dunner zijn). In Hoofdstuk 3 werden impliciete overtuigingen over actueel en ideaal lichaamsbeeld gecapteerd met twee versies van de zgn. Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2006). In Hoofdstuk 4 beschrijf ik de ontwikkeling en validatie van de zgn. Relational Responding Task (RRT). De RRT is een reactietijdtaak die werd ontworpen op basis van de premisse dat propositionele informatie, indien verwerkt, automatisch wordt geëvalueerd als waar of vals op een manier die in overeenstemming is met iemands persoonlijke overtuiging. In Hoofdstuk 4 werd de RRT getest in de context van racistische overtuigingen. Deelnemers werden, in de twee verschillende condities van de RRT, gevraagd om te reageren op verschillende statements (bv. “Vlaamse personen zijn intelligenter dan immigranten") *alsof* ze geloofden dat een bepaald type propositie juist was (bv. dat Vlaamse personen intelligenter zijn dan immigranten) of fout was (bv. dat Vlaamse personen dommer zijn dan immigranten). Het verschil in reactietijd tussen de twee condities (i.e., het RRT effect) bleek te correleren de mate waarin deelnemers expliciet te kennen gaven dat Vlaamse mensen volgens hen intelligenter zijn dan immigranten evenals met vragenlijsten die subtiel en openlijk racisme capteerden. Na deze eerste toepassing van de RRT werd in deze nieuwe procedure ingezet om lichaamsontevredenheid te meten (Hoofdstuk 5). Impliciete overtuigingen omtrent actueel en ideaal lichaamsbeeld werden afzonderlijk gemeten met behulp van twee RRT's. De RRT-effecten toonden aan dat beide overtuigingen van elkaar verschilden en elk een significante en unieke bijdrage leverden aan de voorspelling van lichaamsontevredenheid. Samengevat, in hoofdstukken 3 tot 5 werd bijkomende evidentie verworven voor de hypothese dat het subjectieve waarheidsgehalte van propositionele stellingen kan worden geactiveerd onder automaticiteitscondities. Deze conclusie is moeilijk te verenigen met de centrale assumptie van twee-proces modellen van cognitie dat automatische prikkelverwerking per definitie gebaseerd is op associatieve processen (bv., Rydell & McConnell, 2006). Anderzijds zijn onze bevindingen wel perfect te integreren propositionele één-proces modellen van cognitie. Dit neemt echter niet weg dat de mate waarin propositionele kennis onder automaticiteitscondities kan verwerkt worden allicht aan een aantal randvoorwaarden is gebonden. In de algemene discussie stel ik daarom voor om mentale processen niet langer te beschrijven in functie van de inhoud ervan (bv., propositionele kennis vs. associatieve kennis) maar in functie van de automaticiteitscondities waaronder deze processen kunnen opereren. Hierbij voer ik tevens aan dat de kans dat een bepaald proces kan optreden onder automaticiteitscondities allicht afneemt naarmate de complexiteit ervan toeneemt. In het licht van dit raamwerk worden de RRT-effecten kritisch besproken. Een discussie over de praktische bruikbaarheid van (propositionele) impliciete maten sluit de dissertatie af.

REFERENTIES

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, *32*, 169–177.

Cash, T. F. (1989). Body-image affect: Gestalt versus summing the parts. *Perceptual and Motor Skills*, *69*, 17–18. doi: 10.2466/pms.1989.69.1.17

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*, 1–20. doi: 10.3758/LB.37.1.1

De Houwer, J. (2014). A propositional model of implicit evaluation. *Social and Personality Psychology Compass*, *8*, 342–353. doi: 10.1111/spc3.12111

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238. doi: 10.1037/0022-3514.50.2.229

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692–731. doi: 10.1037/0033-2909.132.5.692

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*, 183–198. doi: 10.1017/S0140525X09000855

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: a systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*, 995–1008. doi: 10.1037/0022-3514.91.6.995

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247. doi: 10.1207/s15327957pspr0803_1

APPENDIX C

DATA STORAGE FACT SHEETS [1]

Chapter 2

```
% Data Storage Fact Sheet

% Name/identifier study: Priming Studies on the Automatic Relational Processing
of Size Relations / internal liplab code: MNH_2012_size-priming (containing 01
- MNH_2012_size_blocked, 02 - MNH_2013_size_neutral, 03 - MNH_2013_size_color,
04 - MNH_2013_size_subliminal)
% Author: Niclas Heider
% Date: 22.4.2015

1. Contact details
============================================================

1a. Main researcher
------------------------------------------------------------
- name: Niclas Heider
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: niclas.heider@ugent.be

1b. Responsible Staff Member (ZAP)
------------------------------------------------------------
- name: Adriaan Spruyt (PhD Promoter)
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: adriaan.spruyt@ugent.be

If a response is not received when using the above contact details, please send
an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology
and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
============================================================
* Reference of the publication in which the datasets are reported:
Heider, N., Spruyt, A., & De Houwer, J. (2015). On the Automaticity of
Relational Stimulus Processing. Manuscript Submitted for Publication.

* Which datasets in that publication does this sheet apply to?:
All datasets reported in the abovementioned publication
All datasets reported in Chapter 2 of the doctoral dissertation of Niclas
Heider

3. Information about the files that have been stored
============================================================

3a. Raw data
------------------------------------------------------------
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

* On which platform are the raw data stored?
  - [X] researcher PC
```

[1] The data storage fact sheets manual (Version March 12th 2014) was used as reference.

```
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another
person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files
-----------------------------------------------------------
* Which other files have been stored?
  - [X] file(s) describing the transition from raw data to reported results.
        R-scripts can be found in subfolder "03Analysis/Calculations"
  - [X] file(s) containing processed data.
        merged raw data and working data can be found in subfolder
        "03Analysis/Results"
  - [X] file(s) containing analyses.
        R-script can be found in subfolder "03Analysis/Calculations"
  - [X] files(s) containing information about informed consent
        Scanned informed consents can be found in subfolder
        "01Background/informed consents" of each study
  - [ ] a file specifying legal and ethical provisions
  - [X] file(s) that describe the content of the stored files and how this
        content should be interpreted.
        Data is stored and organized in line with the guidelines of the Liplab
        (provided by Maarten DeSchryver)
        a readme-file can be found on lowest folder level, specific files are
        in their respective folders
  - [X] other files.
        the computer program can be found in subfolder "01Background"

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [X] other: LipLab Share

* Who has direct access to these other files (i.e., without intervention of
another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
===========================================================
* Have the results been reproduced independently?: [ ] YES / [X] NO

v0.2
```

Chapter 3

```
% Data Storage Fact Sheet

% Name/identifier study: IRAP Study on implicit beliefs about actual and ideal
body image / internal liplab code: MNH_2013_BD_IRAP
% Author: Niclas Heider
% Date: 23.4.2015

1. Contact details
============================================================

1a. Main researcher
------------------------------------------------------------
- name: Niclas Heider
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: niclas.heider@ugent.be

1b. Responsible Staff Member (ZAP)
------------------------------------------------------------
- name: Adriaan Spruyt (PhD Promoter)
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: adriaan.spruyt@ugent.be

If a response is not received when using the above contact details, please send
an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology
and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
============================================================
* Reference of the publication in which the datasets are reported: Heider, N.,
Spruyt, A., & De Houwer, J. (2015). On the Relationship between Body
Dissatisfaction and Implicit Beliefs about Actual and Ideal Body Image.
Manuscript Submitted for Publication.

* Which datasets in that publication does this sheet apply to?:
All datasets reported in the abovementioned publication
All datasets reported in Chapter 3 of the doctoral dissertation of Niclas
Heider

3. Information about the files that have been stored
============================================================

3a. Raw data
------------------------------------------------------------
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

* On which platform are the raw data stored?
  - [X] researcher PC
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another
person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files
------------------------------------------------------------
* Which other files have been stored?
  - [X] file(s) describing the transition from raw data to reported results.
        R-scripts can be found in subfolder "03Analysis/Calculations"
  - [X] file(s) containing processed data.
```

```
        merged raw data and working data can be found in subfolder
        "03Analysis/Results"
  - [X] file(s) containing analyses.
        R-script can be found in subfolder "03Analysis/Calculations"
  - [X] files(s) containing information about informed consent
        Scanned informed consents can be found in subfolder
        "01Background/informed consents" of each study
  - [X] a file specifying legal and ethical provisions
        PDF is stored in subfolder "06EthicalProtocol"
  - [X] file(s) that describe the content of the stored files and how this
        content should be interpreted.
        Data is stored and organized in line with the guidelines of the Liplab
        (provided by Maarten DeSchryver)
        a readme-file can be found on lowest folder level, specific files are
        in their respective folders
  - [X] other files.
        the computer program can be found in subfolder "01Background"

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [X] other: LipLab Share

* Who has direct access to these other files (i.e., without intervention of
another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
===========================================================
* Have the results been reproduced independently?: [ ] YES / [X] NO

v0.2
```

Chapter 4

```
% Data Storage Fact Sheet

% Name/identifier study: RRT study on racism / internal liplab code:
MNH_2014_racial_RRT
% Author: Niclas Heider
% Date: 23.4.2015

1. Contact details
============================================================

1a. Main researcher
------------------------------------------------------------
- name: Niclas Heider
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: niclas.heider@ugent.be

1b. Responsible Staff Member (ZAP)
------------------------------------------------------------
- name: Adriaan Spruyt (PhD Promoter)
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: adriaan.spruyt@ugent.be

If a response is not received when using the above contact details, please send
an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology
and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
============================================================
* Reference of the publication in which the datasets are reported: De Houwer,
J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The Relational
Responding Task: Toward a New Implicit Measure of Beliefs. Frontiers in
Psychology, 6, 319.

* Which datasets in that publication does this sheet apply to?:
All datasets reported in the abovementioned publication
All datasets reported in Chapter 4 of the doctoral dissertation of Niclas
Heider

3. Information about the files that have been stored
============================================================

3a. Raw data
------------------------------------------------------------
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

* On which platform are the raw data stored?
  - [X] researcher PC
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another
person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files
------------------------------------------------------------
* Which other files have been stored?
  - [X] file(s) describing the transition from raw data to reported results.
        R-scripts can be found in subfolder "03Analysis/Calculations"
  - [X] file(s) containing processed data.
```

```
        merged raw data and working data can be found in subfolder
        "03Analysis/Results"
  - [X] file(s) containing analyses.
        R-script can be found in subfolder "03Analysis/Calculations"
  - [X] files(s) containing information about informed consent
        Scanned informed consents can be found in subfolder
        "01Background/informed consents" of each study
  - [ ] a file specifying legal and ethical provisions
  - [X] file(s) that describe the content of the stored files and how this
        content should be interpreted.
        Data is stored and organized in line with the guidelines of the Liplab
        (provided by Maarten DeSchryver)
        a readme-file can be found on lowest folder level, specific files are
        in their respective folders
  - [X] other files.
        the computer program can be found in subfolder "01Background"

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [X] other: LipLab Share

* Who has direct access to these other files (i.e., without intervention of
another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
===========================================================
* Have the results been reproduced independently?: [X] YES / [ ] NO

* If yes, by whom (add if multiple):
   - name: Maarten DeSchryver
   - address: Henri Dunantlaan 2, 9000 Gent
   - affiliation: Ugent
   - e-mail: Maarten.DeSchryver@UGent.be

v0.2
```

Chapter 5

% Data Storage Fact Sheet

% Name/identifier study: RRT study on Body Dissatisfaction / internal liplab code: MNH_2014_BD_RRT
% Author: Niclas Heider
% Date: 23.4.2015

1. Contact details
===

1a. Main researcher

- name: Niclas Heider
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: niclas.heider@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Adriaan Spruyt (PhD Promoter)
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: adriaan.spruyt@ugent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
===
* Reference of the publication in which the datasets are reported: Heider, N., Spruyt, & De Houwer, J.(2015). Body Dissatisfaction Revisited:
The Importance of Beliefs concerning both Actual and Ideal Body Image.
Manuscript Submitted for Publication.

* Which datasets in that publication does this sheet apply to?:
All datasets reported in the abovementioned publication
All datasets reported in Chapter 5 of the doctoral dissertation of Niclas Heider

3. Information about the files that have been stored
===

3a. Raw data
----
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

* On which platform are the raw data stored?
  - [X] researcher PC
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3. Information about the files that have been stored
===

3a. Raw data
----
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

```
* On which platform are the raw data stored?
  - [X] researcher PC
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another
person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files
-----------------------------------------------------------
* Which other files have been stored?
  - [X] file(s) describing the transition from raw data to reported results.
        R-scripts can be found in subfolder "03Analysis/Calculations"
  - [X] file(s) containing processed data.
        merged raw data and working data can be found in subfolder
        "03Analysis/Results"
  - [X] file(s) containing analyses.
        R-script can be found in subfolder "03Analysis/Calculations"
  - [X] files(s) containing information about informed consent
        Scanned informed consents can be found in subfolder
        "01Background/informed consents" of each study
  - [X] a file specifying legal and ethical provisions
        PDF is stored in subfolder "06EthicalProtocol"
  - [X] file(s) that describe the content of the stored files and how this
        content should be interpreted.
        Data is stored and organized in line with the guidelines of the Liplab
        (provided by Maarten DeSchryver)
        a readme-file can be found on lowest folder level, specific files are
        in their respective folders
  - [X] other files.
        the computer program can be found in subfolder "01Background"

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [X] other: LipLab Share

* Who has direct access to these other files (i.e., without intervention of
another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
===========================================================
* Have the results been reproduced independently?: [ ] YES / [X] NO

v0.2
```

Appendix A

```
% Data Storage Fact Sheet

% Name/identifier study: MNH_2011_Priming_Task_Misapplication_colored
% Author: Niclas Heider
% Date: 22.4.2015

1. Contact details
=============================================================

1a. Main researcher
-------------------------------------------------------------
- name: Niclas Heider
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: niclas.heider@ugent.be

1b. Responsible Staff Member (ZAP)
-------------------------------------------------------------
- name: Adriaan Spruyt (PhD Promoter)
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: adriaan.spruyt@ugent.be

If a response is not received when using the above contact details, please send
an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology
and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
=============================================================
* Reference of the publication in which the datasets are reported:
PhD Thesis of Niclas Heider

* Which datasets in that publication does this sheet apply to?:
Experiment 1 reported in Appendix A

3. Information about the files that have been stored
=============================================================

3a. Raw data
-------------------------------------------------------------
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

* On which platform are the raw data stored?
  - [X] researcher PC
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another
person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...
3b. Other files
-------------------------------------------------------------
* Which other files have been stored?
  - [X] file(s) describing the transition from raw data to reported results.
        R-scripts can be found in subfolder "03Analysis/Calculations"
  - [X] file(s) containing processed data.
        merged raw data and working data can be found in subfolder
        "03Analysis/Results"
  - [X] file(s) containing analyses.
        R-script can be found in subfolder "03Analysis/Calculations"
  - [X] files(s) containing information about informed consent
        Scanned informed consents can be found in subfolder
        "01Background/informed consents" of each study
```

```
  - [ ] a file specifying legal and ethical provisions
  - [X] file(s) that describe the content of the stored files and how this
        content should be interpreted.
        Data is stored and organized in line with the guidelines of the Liplab
        (provided by Maarten DeSchryver)
        a readme-file can be found on lowest folder level, specific files are
        in their respective folders
  - [X] other files.
        the computer program can be found in subfolder "01Background"

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [X] other: LipLab Share

* Who has direct access to these other files (i.e., without intervention of
another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
===========================================================
* Have the results been reproduced independently?: [ ] YES / [X] NO

v0.2
```

```
% Data Storage Fact Sheet

% Name/identifier study: Task Misapplication Study 2 / internal liplab code:
MNH_2012_Prime_Task_Misapplication_cued
% Author: Niclas Heider
% Date: 22.4.2015

1. Contact details
===========================================================

1a. Main researcher
-----------------------------------------------------------
- name: Niclas Heider
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: niclas.heider@ugent.be

1b. Responsible Staff Member (ZAP)
-----------------------------------------------------------
- name: Adriaan Spruyt (PhD Promoter)
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: adriaan.spruyt@ugent.be

If a response is not received when using the above contact details, please send
an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology
and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
===========================================================
* Reference of the publication in which the datasets are reported:
PhD Thesis of Niclas Heider

* Which datasets in that publication does this sheet apply to?:
Experiment 2 reported in Appendix A

3. Information about the files that have been stored
===========================================================

3a. Raw data
-----------------------------------------------------------
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

* On which platform are the raw data stored?
  - [X] researcher PC
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another
person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files
-----------------------------------------------------------
* Which other files have been stored?
  - [X] file(s) describing the transition from raw data to reported results.
        R-scripts can be found in subfolder "03Analysis/Calculations"
  - [X] file(s) containing processed data.
        merged raw data and working data can be found in subfolder
        "03Analysis/Results"
  - [X] file(s) containing analyses.
        R-script can be found in subfolder "03Analysis/Calculations"
  - [X] files(s) containing information about informed consent
        Scanned informed consents can be found in subfolder
        "01Background/informed consents" of each study
  - [ ] a file specifying legal and ethical provisions
```

```
  - [X] file(s) that describe the content of the stored files and how this
        content should be interpreted.
        Data is stored and organized in line with the guidelines of the Liplab
        (provided by Maarten DeSchryver)
        a readme-file can be found on lowest folder level, specific files are
        in their respective folders
  - [X] other files.
        the computer program can be found in subfolder "01Background"

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [X] other: LipLab Share

* Who has direct access to these other files (i.e., without intervention of
another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
===========================================================
* Have the results been reproduced independently?: [ ] YES / [X] NO

v0.2
```

```
% Data Storage Fact Sheet

% Name/identifier study: evaluative relational priming study / internal liplab
code: MNH_2013_relative-evaluative-priming
% Author: Niclas Heider
% Date: 22.4.2015

1. Contact details
=============================================================

1a. Main researcher
-------------------------------------------------------------
- name: Niclas Heider
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: niclas.heider@ugent.be

1b. Responsible Staff Member (ZAP)
-------------------------------------------------------------
- name: Adriaan Spruyt (PhD Promoter)
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: adriaan.spruyt@ugent.be

If a response is not received when using the above contact details, please send
an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology
and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
=============================================================
* Reference of the publication in which the datasets are reported:
PhD Thesis of Niclas Heider

* Which datasets in that publication does this sheet apply to?:
Experiment 3 reported in Appendix A

3. Information about the files that have been stored
=============================================================

3a. Raw data
-------------------------------------------------------------
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

* On which platform are the raw data stored?
  - [X] researcher PC
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another
person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files
-------------------------------------------------------------
* Which other files have been stored?
  - [X] file(s) describing the transition from raw data to reported results.
        R-scripts can be found in subfolder "03Analysis/Calculations"
  - [X] file(s) containing processed data.
        merged raw data and working data can be found in subfolder
        "03Analysis/Results"
  - [X] file(s) containing analyses.
        R-script can be found in subfolder "03Analysis/Calculations"
  - [X] files(s) containing information about informed consent
        Scanned informed consents can be found in subfolder
        "01Background/informed consents" of each study
  - [ ] a file specifying legal and ethical provisions
```

```
  - [X] file(s) that describe the content of the stored files and how this
        content should be interpreted.
        Data is stored and organized in line with the guidelines of the Liplab
        (provided by Maarten DeSchryver)
        a readme-file can be found on lowest folder level, specific files are
        in their respective folders
  - [X] other files.
        the computer program can be found in subfolder "01Background"

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [X] other: LipLab Share

* Who has direct access to these other files (i.e., without intervention of
another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
===========================================================
* Have the results been reproduced independently?: [ ] YES / [X] NO

v0.2
```

% Data Storage Fact Sheet

% Name/identifier study: MNH_2011_Priming_AMWANT_1
% Author: Niclas Heider
% Date: 22.4.2015

1. Contact details
===

1a. Main researcher

- name: Niclas Heider
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: niclas.heider@ugent.be

1b. Responsible Staff Member (ZAP)

- name: Adriaan Spruyt (PhD Promoter)
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: adriaan.spruyt@ugent.be

If a response is not received when using the above contact details, please send
an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology
and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
===
* Reference of the publication in which the datasets are reported:
PhD Thesis of Niclas Heider

* Which datasets in that publication does this sheet apply to?:
Experiment 4 reported in Appendix A

3. Information about the files that have been stored
===

3a. Raw data
----
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

* On which platform are the raw data stored?
  - [X] researcher PC
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another
person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files
----
* Which other files have been stored?
  - [X] file(s) describing the transition from raw data to reported results.
        R-scripts can be found in subfolder "03Analysis/Calculations"
  - [X] file(s) containing processed data.
        merged raw data and working data can be found in subfolder
        "03Analysis/Results"
  - [X] file(s) containing analyses.
        R-script can be found in subfolder "03Analysis/Calculations"
  - [X] files(s) containing information about informed consent
        Scanned informed consents can be found in subfolder
        "01Background/informed consents" of each study
  - [ ] a file specifying legal and ethical provisions

```
  - [X] file(s) that describe the content of the stored files and how this
        content should be interpreted.
        Data is stored and organized in line with the guidelines of the Liplab
        (provided by Maarten DeSchryver)
        a readme-file can be found on lowest folder level, specific files are
        in their respective folders
  - [X] other files.
        the computer program can be found in subfolder "01Background"

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [X] other: LipLab Share

* Who has direct access to these other files (i.e., without intervention of
another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
===========================================================
* Have the results been reproduced independently?: [ ] YES / [X] NO

v0.2
```

```
% Data Storage Fact Sheet

% Name/identifier study: MNH_2012_Priming_AMWANT_2
% Author: Niclas Heider
% Date: 22.4.2015

1. Contact details
=============================================================

1a. Main researcher
-------------------------------------------------------------
- name: Niclas Heider
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: niclas.heider@ugent.be

1b. Responsible Staff Member (ZAP)
-------------------------------------------------------------
- name: Adriaan Spruyt (PhD Promoter)
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: adriaan.spruyt@ugent.be

If a response is not received when using the above contact details, please send
an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology
and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
=============================================================
* Reference of the publication in which the datasets are reported:
PhD Thesis of Niclas Heider

* Which datasets in that publication does this sheet apply to?:
Experiment 5 reported in Appendix A

3. Information about the files that have been stored
=============================================================

3a. Raw data
-------------------------------------------------------------
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

* On which platform are the raw data stored?
  - [X] researcher PC
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another
person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files
-------------------------------------------------------------
* Which other files have been stored?
  - [X] file(s) describing the transition from raw data to reported results.
        R-scripts can be found in subfolder "03Analysis/Calculations"
  - [X] file(s) containing processed data.
        merged raw data and working data can be found in subfolder
        "03Analysis/Results"
  - [X] file(s) containing analyses.
        R-script can be found in subfolder "03Analysis/Calculations"
  - [X] files(s) containing information about informed consent
        Scanned informed consents can be found in subfolder
        "01Background/informed consents" of each study
  - [ ] a file specifying legal and ethical provisions
```

```
  - [X] file(s) that describe the content of the stored files and how this
        content should be interpreted.
        Data is stored and organized in line with the guidelines of the Liplab
        (provided by Maarten DeSchryver)
        a readme-file can be found on lowest folder level, specific files are
        in their respective folders
  - [X] other files.
        the computer program can be found in subfolder "01Background"

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [X] other: LipLab Share

* Who has direct access to these other files (i.e., without intervention of
another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
===========================================================
* Have the results been reproduced independently?: [ ] YES / [X] NO

v0.2
```

```
% Data Storage Fact Sheet

% Name/identifier study: relational priming study with body dissatisfied
participants / internal liplab code: MNH_2013_Priming_BD
% Author: Niclas Heider
% Date: 22.4.2015

1. Contact details
=============================================================

1a. Main researcher
-------------------------------------------------------------
- name: Niclas Heider
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: niclas.heider@ugent.be

1b. Responsible Staff Member (ZAP)
-------------------------------------------------------------
- name: Adriaan Spruyt (PhD Promoter)
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: adriaan.spruyt@ugent.be

If a response is not received when using the above contact details, please send
an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology
and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
=============================================================
* Reference of the publication in which the datasets are reported:
PhD Thesis of Niclas Heider

* Which datasets in that publication does this sheet apply to?:
Experiment 6 reported in Appendix A

3. Information about the files that have been stored
=============================================================

3a. Raw data
-------------------------------------------------------------
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

* On which platform are the raw data stored?
  - [X] researcher PC
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another
person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files
-------------------------------------------------------------
* Which other files have been stored?
  - [X] file(s) describing the transition from raw data to reported results.
        R-scripts can be found in subfolder "03Analysis/Calculations"
  - [X] file(s) containing processed data.
        merged raw data and working data can be found in subfolder
        "03Analysis/Results"
  - [X] file(s) containing analyses.
        R-script can be found in subfolder "03Analysis/Calculations"
  - [X] files(s) containing information about informed consent
        Scanned informed consents can be found in subfolder
        "01Background/informed consents" of each study
  - [ ] a file specifying legal and ethical provisions
```

```
  - [X] file(s) that describe the content of the stored files and how this
        content should be interpreted.
        Data is stored and organized in line with the guidelines of the Liplab
        (provided by Maarten DeSchryver)
        a readme-file can be found on lowest folder level, specific files are
        in their respective folders
  - [X] other files.
        the computer program can be found in subfolder "01Background"

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [X] other: LipLab Share

* Who has direct access to these other files (i.e., without intervention of
another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
============================================================
* Have the results been reproduced independently?: [ ] YES / [X] NO

v0.2
```

```
% Data Storage Fact Sheet

% Name/identifier study: relational priming study with body dissatisfied
participants / internal liplab code: MNH_2014_RRT_size
% Author: Niclas Heider
% Date: 22.4.2015

1. Contact details
=============================================================

1a. Main researcher
-------------------------------------------------------------
- name: Niclas Heider
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: niclas.heider@ugent.be

1b. Responsible Staff Member (ZAP)
-------------------------------------------------------------
- name: Adriaan Spruyt (PhD Promoter)
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: adriaan.spruyt@ugent.be

If a response is not received when using the above contact details, please send
an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology
and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
=============================================================
* Reference of the publication in which the datasets are reported:
PhD Thesis of Niclas Heider

* Which datasets in that publication does this sheet apply to?:
Experiment 7 reported in Appendix A

3. Information about the files that have been stored
=============================================================

3a. Raw data
-------------------------------------------------------------
* Have the raw data been stored by the main researcher? [X] YES / [ ] NO

* On which platform are the raw data stored?
  - [X] researcher PC
  - [X] research group file server
  - [X] other (specify): Liblab Share

* Who has direct access to the raw data (i.e., without intervention of another
person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

3b. Other files
-------------------------------------------------------------
* Which other files have been stored?
  - [X] file(s) describing the transition from raw data to reported results.
        R-scripts can be found in subfolder "03Analysis/Calculations"
  - [X] file(s) containing processed data.
        merged raw data and working data can be found in subfolder
        "03Analysis/Results"
  - [X] file(s) containing analyses.
        R-script can be found in subfolder "03Analysis/Calculations"
  - [X] files(s) containing information about informed consent
        Scanned informed consents can be found in subfolder
        "01Background/informed consents" of each study
  - [ ] a file specifying legal and ethical provisions
```

```
  - [X] file(s) that describe the content of the stored files and how this
        content should be interpreted.
        Data is stored and organized in line with the guidelines of the Liplab
        (provided by Maarten DeSchryver)
        a readme-file can be found on lowest folder level, specific files are
        in their respective folders
  - [X] other files.
        the computer program can be found in subfolder "01Background"

* On which platform are these other files stored?
  - [X] individual PC
  - [X] research group file server
  - [X] other: LipLab Share

* Who has direct access to these other files (i.e., without intervention of
another person)?
  - [X] main researcher
  - [X] responsible ZAP
  - [X] all members of the research group
  - [ ] all members of UGent
  - [ ] other (specify): ...

4. Reproduction
===========================================================
* Have the results been reproduced independently?: [ ] YES / [X] NO

v0.2
```